Graduate Texts in Mathematics



David Borthwick

Spectral Theory

Basic Concepts and Applications



Graduate Texts in Mathematics 284

Graduate Texts in Mathematics

Graduate Texts in Mathematics bridge the gap between passive study and creative understanding, offering graduate-level introductions to advanced topics in mathematics. The volumes are carefully written as teaching aids and highlight characteristic features of the theory. Although these books are frequently used as textbooks in graduate courses, they are also suitable for individual study.

More information about this series at http://www.springer.com/series/136

David Borthwick

Spectral Theory

Basic Concepts and Applications

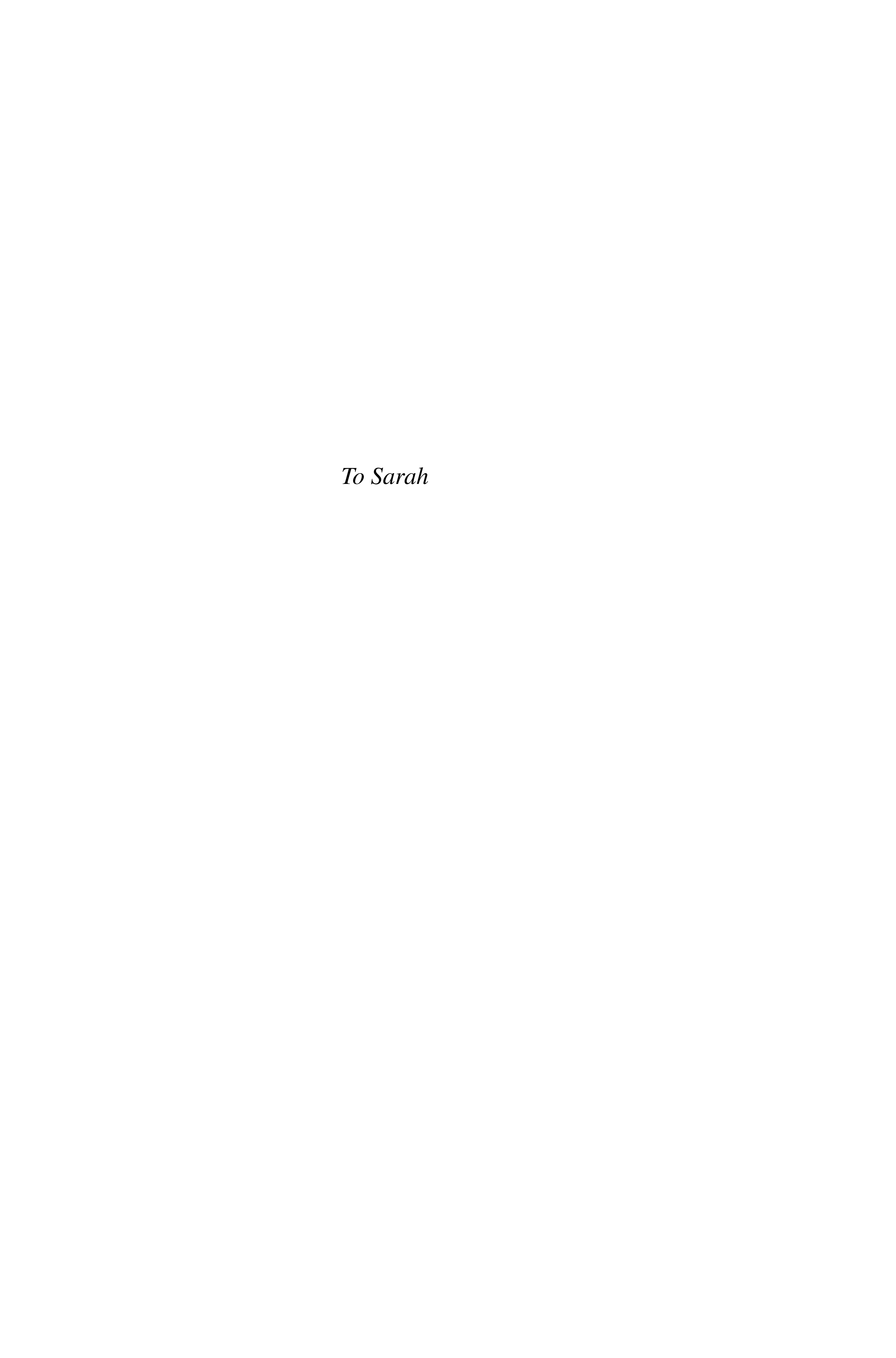


David Borthwick
Department of Mathematics
Emory University
Atlanta, GA, USA

ISSN 0072-5285 ISSN 2197-5612 (electronic)
Graduate Texts in Mathematics
ISBN 978-3-030-38001-4 ISBN 978-3-030-38002-1 (eBook)
https://doi.org/10.1007/978-3-030-38002-1

Mathematics Subject Classification: 35P05, 47A05, 47A10, 46C05, 58J50

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland



To Sarah

Preface

The plan for this book arose from the desire for an introductory text on spectral theory, which would not assume functional analysis as a prerequisite. I wanted this text to include applications involving the Laplacian operator, and yet be concise enough to be covered in a single semester. The inspiration comes in large part from requests for independent study projects from undergraduate or first-year graduate students. Many of these students have proposed topics in particular application areas of spectral theory, such as automorphic forms, differential geometry, or quantum mechanics. Although such applications are covered in sources such as Chavel [19] or Iwaniec [47], books at this level generally assume basic functional analysis and spectral theory as a prerequisite. Most of my independent study students have had some real and complex analysis, but not functional analysis. I wanted to have a text or short course that would bridge the gap between this background and more specialized topics.

Although there are many good introductory books on functional analysis, the shorter, one-semester texts generally do not include enough material on unbounded or differential operators for the applications I had in mind. Books that do cover this part of spectral theory comprehensively, such as the excellent series of Reed and Simon [69–72], are much longer and therefore less suitable for a quick introduction.

The present text thus represents my attempt to produce a short, accessible account of spectral theory that could serve as an introduction to a broad variety of application areas involving the Laplacian operator. It is primarily based on notes from a functional analysis course that I first gave about 15 years ago. To fit both the introductory theory and some interesting applications into one semester posed a significant challenge; it is already difficult to squeeze the essentials of functional analysis into a single term. My strategy was to focus on a relatively small list of applications (Weyl's law for bounded domains, the theory of Schrödinger operators with positive potentials, etc.). I built the first half of the course around these topics, limiting the coverage of functional analysis background to the material that was necessary for the chosen examples.

The outline for this book was developed by the same approach. The result is a treatment of functional analysis that differs from more traditional texts in two major

ways. First, the focus is almost exclusively on separable Hilbert spaces, and much of the Banach space theory is omitted. Second, the theory of unbounded operators is developed from the beginning, rather than as an addendum to the bounded case. Applications to differential operators are introduced as early as possible, mainly in the examples.

After a brief historical introduction in Chapter 1, the main body of the text is roughly divided into two parts. Chapters 2 through 5 cover the theoretical background, from the theory of Hilbert spaces and unbounded operators to the proof of the spectral theorem. These chapters are sequential and strongly interdependent. The second part, consisting of Chapters 6 through 9, is devoted to more specific contexts such as the Dirichlet Laplacian or Schrödinger operators. These later chapters are essentially independent and could be read in any order. This structure provides the flexibility to support a one-semester functional course, a special topics course, or an independent study project in a particular application area.

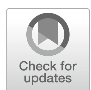
The book is aimed at an advanced undergraduate or beginning graduate level. The reader is assumed to have background including real and complex analysis, measure theory, and linear algebra, but no previous knowledge of functional analysis. The necessary background material is sketched in the appendix, with references.

Acknowledgments It is a pleasure to thank Evans Harrell for encouragement and advice regarding the content of this text. I am also grateful to my students Kenny Jones and Varoon Pazhyanur, who worked through parts of the manuscript while it was in development. Thanks also to Loretta Bartolini at Springer, for patience and helpful advice during the development process, and to the series editors and anonymous reviewers for helpful comments and suggestions.

Atlanta, GA, USA David Borthwick
October 2019

Contents

Chapter 1
Introduction



In 1671 Isaac Newton adopted the Latin word "spectrum" to describe the spread of colors illustrated in Figure 1.1, produced by passing white light through a prism. The origins of spectral theory can be traced much earlier, however, in the context of sound waves. Pythagoras discovered the relationship between the length of a vibrating string and the musical pitch it produces, some 2000 years before Newton. This observation is arguably the starting point for spectral theory.

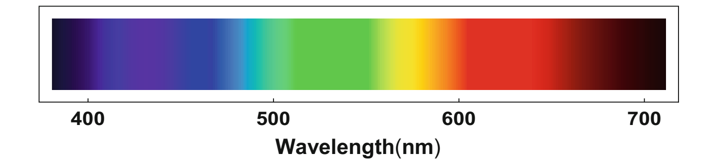


Fig. 1.1 The spectrum of visible light

Marin Mersenne refined the work of Pythagoras in 1636, noting that the vibrating string produces, in addition to its fundamental tone, a simultaneous set of overtones. The frequencies of these overtones, visible as sharp peaks in Figure 1.2, are integer multiples of the fundamental frequency. A century after Mersenne, Jean-Baptiste d'Alembert explained the overtone phenomenon by developing the wave equation as a model for the motion of the string. The wave equation admits sinusoidal solutions whose frequencies come in integer multiples. This is the first historical case of an observable spectral phenomenon being explained in terms of differential equations.

Another great milestone in the mathematical development of spectral theory is Joseph Fourier's study of the heat equation in 1807, based on the decomposition of functions into trigonometric series. Fourier's work led to the first systematic treatment of the spectra of ordinary differential equations by Charles Sturm and Joseph Liouville, in the 1830s. The spectral theory of partial differential operators was subsequently developed by Gustav Dirichlet and Henri Poincaré, among others.

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1_1

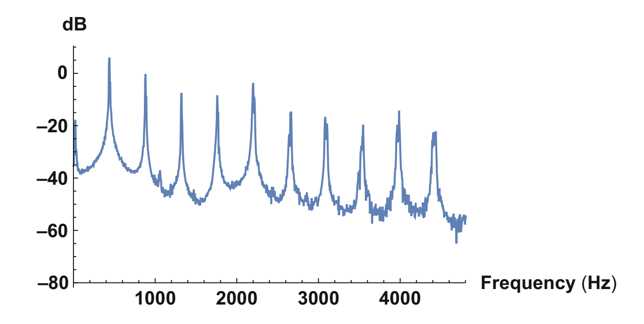


Fig. 1.2 The frequency spectrum of a violin A string, tuned to 440 Hz

Matrix theory was also emerging during the same period. The concept of the spectrum of a matrix first appears in the late eighteenth century, in the work of Joseph-Louis Lagrange. Lagrange defined the moments of inertia of a rigid body in terms of the characteristic values of a matrix. This application inspired Augustin-Louis Cauchy to prove the spectral theorem for real symmetric matrices in 1826, which was generalized to complex self-adjoint matrices by Charles Hermite in 1855.

With hindsight, the parallels between the matrix and differential operator versions of spectral theory are evident. The connection did not become explicit, however, until the early 1900s. David Hilbert, building on the work of Vito Volterra and Ivar Fredholm on integral equations, developed the spectral theory of integral operators. This essentially generalizes the matrix theory to the context of infinite-dimensional function spaces. The link to differential operators lies in the fact that the primary integral operators of interest were solution operators (Green's functions) for classical PDE. Hilbert coined the term "eigenvalue" and was the first to refer to the set of eigenvalues of a matrix or operator as its spectrum. His research laid the foundation for the modern development of spectral theory and functional analysis.

On the empirical side, physical observations of the spectrum had evolved dramatically in the nineteenth century. In 1802 William Wollaston introduced a spectrometer accurate enough to detect dark absorption lines in the spectrum of sunlight. Anders Jonas Ångström discovered that hot gases produce characteristic emission lines around 1860, and was able to deduce the presence of hydrogen in the sun by observing the characteristic emission lines shown in Figure 1.3. A few years later, Jules Janssen discovered a new element, helium, by analyzing the solar spectrum during an eclipse.

The existence of these absorption and emission lines posed a serious challenge for classical physics: Why would atomic spectra be discrete? This challenge was soon compounded by other physical observations, involving black body radiation and the photoelectric effect, which revealed spectral anomalies not explained by classical models. The need to explain these discrepancies ultimately led to the development of quantum mechanics in the early twentieth century, by pioneers including Max Planck, Niels Bohr, Werner Heisenberg, Louis de Broglie, Arthur Compton, Albert Einstein, Erwin Schrödinger, and Max Born. The turning point for the speculative quantum theory came in 1926, when Schrödinger published his



Fig. 1.3 The visible hydrogen emission lines in the Balmer series

analysis of the spectrum of atomic hydrogen. Schrödinger showed that the atomic emission lines were determined by the spectrum of a particular differential operator, as predicted by the quantum theory.

Inspired by these revolutionary developments in physics, Marshall Stone and John von Neumann sought to provide a mathematical foundation for quantum mechanics by extending the spectral theory of Hilbert. Independently, during the period 1929–1932, they established a general spectral theorem for linear operators. This result generalizes the matrix theory of Cauchy and Hermite to a context that includes partial differential operators, finally pulling together the distinct lines of the development of spectral theory into a consistent framework.

The advances by Stone and von Neumann set the stage for the rapid development of spectral theory as a research area in mathematics. Although quantum mechanics provided much of the initial motivation, the modern development of the field is notable for rich connections to other areas of mathematics, including partial differential equations, harmonic analysis, number theory, differential geometry, and operator algebras.

Notes

This brief historical sketch emphasizes topics covered later in this book and is not meant to be comprehensive. Sources for this material include Birkhoff and Kreysig [12], Dieudonné [25], Monna [63], and Steen [86].

Chapter 2
Hilbert Spaces



To set the stage for our development of spectral theory, we must first introduce some basic tools of functional analysis. The scope of this discussion will be limited to those aspects of the theory which will be needed later in the book. This means that the focus is primarily on Hilbert spaces, and we will sometimes state results in less than full generality for the sake of simplified exposition.

2.1 Normed Vector Spaces

Vector spaces in this book are assumed to be defined over the complex numbers, unless otherwise indicated. The standard finite-dimensional example is $\mathbb{C}^n$. The length of a vector $z \in \mathbb{C}^n$ is defined by

$$\begin{aligned} |z| &:= \sqrt{\bar{z} \cdot z} \\ &= \sqrt{|z_1|^2 + \cdots + |z_n|^2}, \end{aligned} \tag{2.1}$$

where $z = (z_1, \ldots, z_n)$. Note that $|z|$ is equal to the Euclidean length of the corresponding vector in $\mathbb{R}^{2n}$. This length function serves as the prototype for the following:

Definition 2.1. A *norm* on a complex vector space $\mathcal{V}$ is a function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ satisfying, for all $v, w \in \mathcal{V}$ and $a \in \mathbb{C}$:

(i) positive definiteness: $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$;
(ii) homogeneity: $\|av\| = |a|\|v\|$;
(iii) triangle inequality: $\|v + w\| \leq \|v\| + \|w\|$.

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1_2

Example 2.2. For a compact subset $K \subset \mathbb{R}^n$, let $C(K)$ denote the space of continuous functions $K \to \mathbb{C}$. Continuous functions on a compact set are bounded, so a natural choice of norm is

$$\|f\| := \sup_{x \in K} |f(x)|.$$

It is easy to check that the properties of Definition 2.1 are satisfied in this case. ◊

A normed vector space $\mathcal{V}$ is naturally equipped with a metric topology defined by the distance function

$$\operatorname{dist}(v, w) := \|v - w\|.$$

In particular, a sequence $\{v_n\} \subset \mathcal{V}$ converges to $w \in \mathcal{V}$ if

$$\lim_{n \to \infty} \|v_n - w\| = 0.$$

Because we are often trying to establish the existence of a limit, it is extremely useful to have a notion of convergence that does not make reference to the limit vector. The sequence $\{v_n\}$ is *Cauchy* if

$$\lim_{n,m \to \infty} \|v_n - v_m\| = 0.$$

A convergent sequence is automatically Cauchy, by the triangle inequality. If every Cauchy sequence in $\mathcal{V}$ converges in $\mathcal{V}$, then $\mathcal{V}$ is said to be *complete* as a metric space. Such spaces were studied extensively by Stefan Banach in the early 1920s.

Definition 2.3. A complete normed vector space is called a *Banach space*.

Series are also well defined in a normed vector space, with convergence defined in terms of the limit of partial sums. We say that a series $\sum u_k$ with $u_k \in \mathcal{V}$ is *absolutely convergent* if

$$\sum_{k=1}^{\infty} \|u_k\| < \infty.$$

The completeness of a normed vector space can also be formulated in terms of the connection between convergence and absolute convergence.

Theorem 2.4. *A normed vector space $\mathcal{V}$ is complete if and only if every absolutely convergent series is convergent.*

Proof Suppose that $\mathcal{V}$ is complete, and let $\sum u_k$ be an absolutely convergent series. The difference of two partial sums can be estimated by the triangle inequality,

$$\left\| \sum_{k=1}^{m} u_k - \sum_{k=1}^{n} u_k \right\| = \left\| \sum_{k=n+1}^{m} u_k \right\|$$
$$\leq \sum_{k=n+1}^{m} \|u_k\|,$$

assuming $m > n$. Since $\sum \|u_k\| < \infty$, this shows that the sequence of partial sums is Cauchy. Therefore, the series $\sum u_k$ converges in $\mathcal{V}$ by the completeness hypothesis.

Now assume that all absolutely convergent sequences converge in $\mathcal{V}$, and let $\{w_n\}$ be a Cauchy sequence. Using the Cauchy property, we can choose a subsequence $\{w_{n_k}\}$ such that

$$\|w_{n_k} - w_{n_{k+1}}\| \leq 2^{-k} \tag{2.2}$$

for $k \in \mathbb{N}$. Let $u_1 := w_{n_1}$ and $u_k := w_{n_k} - w_{n_{k-1}}$ for $k \geq 2$, so that

$$\sum_{k=1}^{m} u_k = w_{n_m}. \tag{2.3}$$

By (2.2), the series $\sum u_k$ converges absolutely, and therefore the subsequence $\{w_{n_m}\}$ converges. A Cauchy sequence with a convergent subsequence is convergent, by the triangle inequality. □

2.2 L^p Spaces

A *measure space* is a triple $(X, \mathcal{M}, \mu)$ consisting of a set X, a σ-algebra $\mathcal{M}$ of subsets of X, and a countably additive measure function $\mu : \mathcal{M} \to [0, \infty]$. We will generally assume that X is σ-finite, meaning that X admits decomposition into a countable union of sets of finite measure. Associated with the measure is an integral denoted by

$$f \mapsto \int_X f \, d\mu,$$

defined for a measurable function f as long as $f \geq 0$ or $\int_X |f| < \infty$. (See Appendix A.1 for a brief review of measure and integration theory.) In the case of Lebesgue measure on $\mathbb{R}^n$, we resort to the usual calculus notation, with $d^n x$ denoting the Lebesgue integral.

To each measure space we associate a family of normed function spaces $L^p(X, d\mu)$, for $1 \leq p \leq \infty$. For finite p this consists of measurable functions for which

$$\int_X |f|^p \, d\mu < \infty, \tag{2.4}$$

subject to the equivalence relation $f \sim g$ if $f = g$ almost everywhere with respect to μ. For a subset $\Omega \subset \mathbb{R}^n$, Lebesgue measure is assumed by default, and we simply write $L^p(\Omega)$ in this case.

The condition (2.4) defines a linear space by the convexity of the function x^p for $p \in [1, \infty)$. The natural choice for a norm on $L^p(X, d\mu)$ is

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}.$$

The power $1/p$ is included for the sake of homogeneity, and the equivalence relation guarantees that $\|\cdot\|_p$ is positive definite. The triangle inequality for $\|\cdot\|_p$ is known as the *Minkowski inequality*:

$$\|f + g\|_p \le \|f\|_p + \|g\|_p. \tag{2.5}$$

See Appendix A.2 for the proof.

For a continuous, compactly supported function f, it is easy to check that

$$\lim_{p \to \infty} \|f\|_p = \sup |f|.$$

This suggests that the appropriate extension of the L^p norm to the case $p = \infty$ should be a sup norm. We therefore define $L^\infty(X, d\mu)$ as the space of functions $f : X \to \mathbb{C}$ which are bounded almost everywhere with respect to μ, under the same equivalence relation as for the other L^p spaces. The sup norm is adapted to the equivalence relation by setting

$$\|f\|_\infty := \inf\left\{m \in \mathbb{R} : \ |f| \le m \text{ a.e.}\right\}.$$

This is called the *essential supremum* of $|f|$.

We use a lowercase ℓ to distinguish L^p spaces based on counting measure. That is, if X is a discrete set, such as $\mathbb{N}$ or $\mathbb{Z}$, then

$$\ell^p(X) := L^p(X, d\nu),$$

where ν denotes counting measure. For example, functions $\mathbb{N} \to \mathbb{C}$ are naturally identified with sequences, and integration with respect to counting measure is series summation. Thus, the norm on $\ell^p(\mathbb{N})$ is given by

$$\|(a_1, a_2, \dots)\|_p := \left(\sum_{j=1}^{\infty} |a_j|^p \right)^{\frac{1}{p}}.$$

One of the fundamental results of the Lebesgue integration theory is the completeness of L^p spaces. This was originally proven independently for $p = 2$ by Frigyes Riesz and Ernst Fischer.

Theorem 2.5 (Riesz–Fischer). *For $p \in [1, \infty]$, $L^p(X, d\mu)$ is a Banach space.*

The proof is given in Appendix A.2.

2.3 Bounded Linear Maps

A linear map between vector spaces is called an *operator*. We will adopt a more specialized usage of this term in Section 3.1, in the context of Hilbert spaces. Here we focus on more basic definitions and results. Throughout this section, $\mathcal{V}$ and $\mathcal{W}$ are assumed to be normed vector spaces, not necessarily complete unless explicitly indicated.

Definition 2.6. An operator $T : \mathcal{V} \to \mathcal{W}$ is *bounded* if there exists a constant $C > 0$ such that

$$\|Tv\| \leq C\|v\|, \quad \text{for all } v \in \mathcal{V}. \tag{2.6}$$

The space of bounded operators $\mathcal{V} \to \mathcal{W}$ is denoted by $\mathcal{L}(\mathcal{V}, \mathcal{W})$, which is simplified to $\mathcal{L}(\mathcal{V})$ in the case $\mathcal{W} = \mathcal{V}$.

It is not difficult to see that a linear map satisfies (2.6) if and only if it is continuous; the proof is left to Exercise 2.1. The set $\mathcal{L}(\mathcal{V}, \mathcal{W})$ clearly forms a vector space. The *operator norm*, defined for $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ by

$$\|T\| := \sup_{\|v\|=1} \|Tv\| \tag{2.7}$$

is the optimal constant for the inequality (2.6). It is easy to verify that this has the properties of a norm.

The operator norm satisfies a convenient multiplicative estimate: for $T \in \mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ and $S \in \mathcal{L}(\mathcal{V}_2, \mathcal{V}_3)$, we have $ST \in \mathcal{L}(\mathcal{V}_1, \mathcal{V}_3)$ and

$$\|ST\| \leq \|S\|\|T\|. \tag{2.8}$$

This follows directly from the definition (2.7).

Example 2.7. Suppose $\mathcal{V} = \mathbb{C}^n$ with the Euclidean norm (2.1). Then for $T \in \mathcal{L}(\mathbb{C}^n, \mathcal{W})$,

$$\|Tz\| \leq \sum_{j=1}^{n} |z_j| \|Te_j\|,$$

where $z = (z_1, \ldots, z_n)$ and $\{e_j\}$ is the standard coordinate basis. If $M := \max\{\|Te_j\|\}$, then

$$\|Tz\| \le Mn|z|,$$

and so T is bounded. This argument can be extended to show that an operator $\mathcal{V} \to \mathcal{W}$ is bounded whenever $\mathcal{V}$ is finite-dimensional. ♢

Example 2.8. Let $(X, \mathcal{M}, \mu)$ be a measure space. For $f \in L^\infty(X, d\mu)$, we define the *multiplication operator* M_f on $L^p(X, d\mu)$,

$$M_f : v \mapsto fv.$$

Since $|f| \le \|f\|_\infty$ almost everywhere, we can estimate

$$\|fv\|_p \le \|f\|_\infty \|v\|_p.$$

Thus, M_f is bounded with $\|M_f\| \le \|f\|_\infty$.

In fact, these norms are equal. For $a < \|f\|_\infty$, set $A := \{|f| \ge a\}$ and let χ_A denote the characteristic function of A. Note that $\|\chi_A\|_p = \mu(A)$, which is strictly positive by the choice of a. The estimate

$$\|f\chi_A\|_p \ge a\|\chi_A\|_p$$

thus shows that $\|M_f\| \ge a$ for all $a < \|f\|_\infty$. Therefore

$$\|M_f\| = \|f\|_\infty.$$

♢

Example 2.9. Let $\mathcal{V} = C^1[0, 1]$, the space of continuously differentiable functions on the interval $[0, 1]$, and $\mathcal{W} = C[0, 1]$, both equipped with the sup norm. The derivative $T := d/dx$ is well defined as an operator $\mathcal{V} \to \mathcal{W}$, but clearly not bounded. For example, the functions $f_n(x) := x^n$ satisfy $\sup|f_n| = 1$ and $\sup|Tf_n| = n$. ♢

Certain aspects of linear algebra carry over directly to the theory of bounded operators. The *kernel* of $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ is defined by

$$\ker(T) := \{v \in V : Tv = 0\},$$

and the *range* is the image,

$$\mathrm{range}(T) := \{Tv : v \in V\}.$$

The *rank* of T is the dimension of the range, which might be infinite. There is no rank-nullity theorem for T unless $\mathcal{V}$ is finite-dimensional.

By linearity, $\ker(T)$ is a subspace of $\mathcal{V}$, and T is injective if and only if $\ker(T) = \{0\}$. The bijectivity of T is equivalent to the existence of an inverse linear map $T^{-1} : \mathcal{W} \to \mathcal{V}$. However, the inverse map T^{-1} is not necessarily bounded.

An *isometry* is a map between metric spaces that preserves the distance between points. For operators on normed vector spaces, this is equivalent to preserving the norm. That is, an operator $T \in \mathcal{L}(V, W)$ is an isometry if and only if

$$\|Tv\| = \|v\|$$

for all $v \in \mathcal{V}$. Note that isometries are not required to be invertible. An example of this is the right shift operator on $\ell^2(\mathbb{N})$, which maps the sequence $(a_1, a_2, \dots)$ to $(0, a_1, a_2, \dots)$.

2.3.1 Operator Topologies

The metric topology on $\mathcal{L}(V, W)$ defined by the operator norm is naturally called the *operator topology*. The following result shows that this choice is natural, in terms of completeness.

Theorem 2.10. *If $\mathcal{V}$ and $\mathcal{W}$ are normed vector spaces and $\mathcal{W}$ is complete, then $\mathcal{L}(\mathcal{V}, \mathcal{W})$ is complete with respect to the operator norm.*

Proof Let $\{T_n\}$ be a Cauchy sequence in $\mathcal{L}(\mathcal{V}, \mathcal{W})$. For $v \in \mathcal{V}$,

$$\|T_n v - T_m v\| \le \|T_n - T_m\| \|v\|$$

implying that the sequence $\{T_n v\}$ is Cauchy in $\mathcal{W}$. Therefore, by the completeness of $\mathcal{W}$, we can define Tv as the limit

$$Tv := \lim_{n \to \infty} T_n v. \tag{2.9}$$

It follows from the linearity of the maps T_n that T is linear. To see that T is bounded, we note that a Cauchy sequence is necessarily bounded, so there exists a constant $M < \infty$ such that $\|T_n\| \le M$ for all n. It follows that that $\|T\| \le M$ also.

It remains to show that $T_n \to T$ with respect to the operator norm. Since $\{T_n\}$ is Cauchy, given $\varepsilon > 0$ there exists N so that $\|T_n - T_m\| < \varepsilon$ for $n, m \ge N$. For a unit vector v, we can choose $m \ge N$ so that $\|T_m v - Tv\| < \varepsilon$, by the definition (2.9). For $n \ge N$, this gives

$$\begin{aligned} \|(T_n - T)v\| &\le \|(T_n - T_m)v\| + \|T_m v - Tv\| \\ &< 2\varepsilon. \end{aligned}$$

Since N was chosen independently of v, this shows $\|T_n - T\| \to 0$. $\square$

Convergence in operator norm is the analog of uniform convergence for ordinary functions. In fact the operator topology is sometimes referred to as the "uniform" topology.

There are also weaker notions of convergence for operators that prove to be quite useful. For example, the analog of pointwise convergence of functions is the following. We say that $T_n \to T$ in the *strong operator* sense if

$$\lim_{n\to\infty} \|T_n v - Tv\| = 0, \tag{2.10}$$

for all $v \in \mathcal{V}$. Strong operator convergence has already played a role in the proof of Theorem 2.10. The limiting operator T was first obtained as a strong limit in (2.9). We then used the Cauchy assumption to upgrade to operator convergence.

The use of the term "strong" here is perhaps confusing, since strong operator convergence is evidently weaker than operator-norm convergence. The intended comparison is to another topology defined as follows. The *dual* space of a topological vector space $\mathcal{W}$ is

$$\mathcal{W}' := \mathcal{L}(\mathcal{W}, \mathbb{C}),$$

i.e., $\mathcal{W}'$ is the space of continuous linear functionals $\mathcal{W} \to \mathbb{C}$. We say that $T_n \to T$ in the *weak operator* sense if

$$\lim_{n\to\infty} F(T_n v) = F(Tv) \tag{2.11}$$

for all $v \in \mathcal{V}$ and $F \in \mathcal{W}'$. Clearly, strong operator convergence implies weak operator convergence, because of the continuity requirement on functionals in the dual space.

Example 2.11. Suppose T denotes the left shift operator on $\ell^2(\mathbb{N})$, defined by

$$T(a_1, a_2, \ldots) := (a_2, a_3, \ldots).$$

Clearly $\|T^k\| = 1$ for all $k \in \mathbb{N}$, so T^k does not converge to 0 in the operator topology as $k \to \infty$. On the other hand, since

$$\left\|T^k(a_1, a_2, \ldots)\right\|^2 = \sum_{j=k+1}^{\infty} |a_j|^2$$

and $\sum |a_j|^2 < \infty$, we have $T^k \to 0$ in the strong sense.

Now let S denote the right shift operator

$$S(a_1, a_2, \ldots) := (0, a_1, a_2, \ldots).$$

Since $\|S^k a\| = \|a\|$ for each $a \in \ell^2(\mathbb{N})$, S^k clearly does not converge to 0 in the strong operator sense. However, we do have $S^k \to 0$ in the weak operator sense. The Riesz lemma, to be proven in Section 2.6, identifies the dual space of $\ell^2(\mathbb{N})$ with $\ell^2(\mathbb{N})$ itself. That is, a continuous linear functional $F : \ell^2(\mathbb{N}) \to \mathbb{C}$ corresponds to a unique element $b \in \ell^2(\mathbb{N})$ such that

$$F(a) = \langle b, a \rangle.$$

For $k \in \mathbb{N}$,

$$\begin{aligned} \langle b, S^k a \rangle &= \sum_{j=1}^{\infty} \overline{b_{j+k}} a_j \\ &= \langle T^k b, a \rangle. \end{aligned}$$

Since $T^k \to 0$ in the strong operator sense, $S^k \to 0$ in the weak sense. ♢

Example 2.12. Consider multiplication operators on $L^p(X, d\mu)$, defined as in Example 2.8. For a sequence $\{f_n\} \subset L^\infty(X, d\mu)$, assume that $f_n \to 0$ pointwise and $\|f_n\|_\infty \le C$ for some constant C. The dominated convergence theorem implies that

$$\lim_{n \to \infty} \int_X |f_n u|^p \, d\mu = 0$$

for $u \in L^p(X, d\mu)$. Hence, $M_{f_n} \to 0$ in the strong operator sense.

To check for operator-norm convergence, we can use the estimate

$$\|f_n u\|_p \le \|f_n\|_\infty \|u\|_p.$$

This shows that $M_{f_n} \to 0$ in the operator topology provided that $\|f_n\|_\infty \to 0$, i.e., if $f_n \to 0$ uniformly. ♢

2.3.2 Uniform Boundedness

We conclude this section with a fundamental result in functional analysis, sometimes called the Banach–Steinhaus theorem. Our main use for this will be to convert weaker pointwise bounds on operators into estimates of the operator norms.

Theorem 2.13 (Uniform Boundedness Principle). *Let $\mathcal{V}$ be a Banach space and $\mathcal{W}$ a normed vector space. Assume that a subset $\mathcal{F} \subset \mathcal{L}(\mathcal{V}, \mathcal{W})$ satisfies*

$$\sup_{T \in \mathcal{F}} \|Tv\| < \infty$$

for each $v \in \mathcal{V}$*. Then* $\mathcal{F}$ *is uniformly bounded in the sense that*

$$\sup_{T \in \mathcal{F}} \|T\| < \infty.$$

Proof We will prove the contrapositive statement, starting from the assumption that

$$\sup_{T \in \mathcal{F}} \|T\| = \infty.$$

This assumption implies that there exists a sequence $\{T_n\} \subset \mathcal{F}$ such that

$$\|T_n\| \geq 4^n,$$

for all n. Our goal is to construct a convergent sequence $v_n \to v$ such that $\|T_n v_n\|$ grows fast enough to ensure that $\|T_n v\| \to \infty$. This will produce a vector v for which

$$\sup_{T \in \mathcal{F}} \|Tv\| = \infty.$$

To set up the construction of $\{v_n\}$, we first observe that for all $v \in \mathcal{V}$ and $\varepsilon > 0$,

$$\|T_n\| \leq \frac{1}{\varepsilon} \sup_{u \in \mathcal{V}: \|u-v\| < \varepsilon} \|T_n u\|. \tag{2.12}$$

This follows by first noting that the triangle inequality implies

$$\|T_n w\| \leq \frac{1}{2} \big\|T_n(v_0 + w)\big\| + \frac{1}{2} \big\|T_n(v_0 - w)\big\|,$$

and then taking the supremum over $\|w\| < \varepsilon$ on both sides.

Fix a constant c with $\frac{1}{2} < c < 1$. The sequence $\{v_n\}$ is constructed by induction starting from some arbitrary $v_0 \in \mathcal{V}$. Given v_{n-1}, there exists according to (2.12) a vector $v_n \in \mathcal{V}$ such that

$$\|v_n - v_{n-1}\| < 3^{-n}$$

and

$$\|T_n v_n\| \geq c 3^{-n} \|T_n\|. \tag{2.13}$$

The resulting sequence $\{v_n\}$ is Cauchy, because for $m \geq n$,

$$\|v_m - v_n\| \leq \sum_{k=n+1}^{\infty} 3^{-k} = \frac{3^{-n}}{2}. \tag{2.14}$$

By completeness, there exists $v \in \mathcal{V}$ such that $v_n \to v$. Taking $m \to \infty$ in (2.14) gives

$$\|v - v_n\| \leq \frac{3^{-n}}{2}.$$

Combining this with (2.13) yields the lower bound

$$\begin{aligned} \|T_n v\| &\geq \|T_n v_n\| - \|T_n(v - v_n)\| \\ &\geq \left(c - \tfrac{1}{2}\right)3^{-n}\|T_n\|. \end{aligned}$$

Therefore, by the assumption that $\|T_n\| \geq 4^n$,

$$\|T_n v\| \geq \left(c - \tfrac{1}{2}\right)(\tfrac{4}{3})^n.$$

Because $c > \frac{1}{2}$, this yields $\|T_n v\| \to \infty$. □

2.4 Hilbert Spaces

In Euclidean $\mathbb{C}^n$, the square of the norm can be written as a dot product $\bar{z} \cdot z$, which is the restriction to the diagonal of the pairing,

$$(z, w) \mapsto \bar{z} \cdot w.$$

This pairing is called a *sesquilinear* function of (z, w), meaning conjugate-linear in the first variable and linear in the second. The placement of the conjugate is a matter of convention. We follow the quantum mechanics style here in conjugating on the first variable rather than the second.

Definition 2.14. An *inner product* on a complex vector space $\mathcal{V}$ is a sesquilinear pairing $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$ satisfying the following conditions:

(i) $\langle v, v \rangle \geq 0$ for all $v \in \mathcal{V}$, and $\langle v, v \rangle = 0$ if and only if $v = 0$.
(ii) $\langle u, v \rangle = \overline{\langle v, u \rangle}$ for all $u, v \in \mathcal{V}$.

By analogy to the Euclidean case, we can create a norm from the inner product by setting

$$\|v\| := \sqrt{\langle v, v \rangle}. \tag{2.15}$$

Positive definiteness and homogeneity of the norm follow immediately from the defining conditions of the inner product. To establish the triangle inequality, we first prove the following:

Lemma 2.15 (Cauchy–Schwarz Inequality). *If $\langle \cdot, \cdot \rangle$ is an inner product on a complex vector space $\mathcal{V}$, then*

$$|\langle v, w\rangle| \le \|v\| \|w\| \tag{2.16}$$

for $v, w \in \mathcal{V}$, where $\|\cdot\|$ is defined by (2.15).

Proof The result is trivial if $w = 0$, and for $w \neq 0$ we can reduce to the case $\|w\| = 1$ by rescaling. For $\|w\| = 1$ the inequality follows from

$$\begin{aligned} 0 &\le \left\| v - \langle w, v\rangle w \right\|^2 \\ &= \|v\|^2 - |\langle w, v\rangle|^2. \end{aligned}$$

□

Using Cauchy–Schwarz, it is now easy to see the (2.15) satisfies the triangle inequality. We can simply estimate

$$\begin{aligned} \|v + w\|^2 &= \|v\|^2 + 2 \operatorname{Re}\langle v, w\rangle + \|w\|^2 \\ &\le \|v\|^2 + 2\|v\| \|w\| + \|w\|^2 \\ &= (\|v\| + \|w\|)^2. \end{aligned}$$

This completes the argument that (2.15) is a norm.

Example 2.16. For a measure space $(X, \mathcal{M}, \mu)$, consider the space $L^2(X, d\mu)$. The pairing

$$\langle f, g\rangle := \int_X \overline{f} g \, d\mu$$

is related to the L^2 norm by (2.15), and clearly satisfies the properties of an inner product. ◊

None of the other L^p spaces with $p \neq 2$ are inner product spaces. One way to see this is by means of a simple identity from Euclidean geometry.

Lemma 2.17 (Parallelogram Law). *For a normed vector space $\mathcal{V}$, there exists an inner product for which $\|v\|^2 := \langle v, v\rangle$ if and only if*

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$$

for all $u, v \in \mathcal{V}$.

The proof of Lemma 2.17 is a straightforward calculation. The main point is the *polarization identity*,

$$\langle u, v \rangle = \frac{1}{4}\Big(\|u+v\|^2 - \|u-v\|^2 + i\|u+iv\|^2 - i\|u-iv\|^2\Big), \tag{2.17}$$

which is a direct consequence of (2.15). For any norm, the right side of (2.17) defines a candidate for an inner product. The condition that this function is sesquilinear is equivalent to the parallelogram law.

An inner-product space carries a metric topology defined by the canonical norm (2.15). As with normed vector spaces, completeness is an essential property for many applications.

Definition 2.18. A *Hilbert space* $\mathcal{H}$ is a complete inner-product space.

Function spaces with complete inner products played an important role in the work of David Hilbert and others on integral equations, in the first decade of the twentieth century. It was von Neumann who later coined the term "Hilbert space." The fundamental example is $L^2(X, d\mu)$. All of the explicit Hilbert spaces encountered in this book will be either L^2 spaces or related spaces based on some modification of the L^2 bracket.

The polarization identity shows that an isometry of inner product spaces preserves the inner product as well as the norm. A *unitary* map is a bijective isometry between Hilbert spaces. Operators $T_1 \in \mathcal{L}(\mathcal{H}_1)$ and $T_2 \in \mathcal{L}(\mathcal{H}_2)$ are said to be *unitarily equivalent* if they are conjugate to each other by a unitary map $U : \mathcal{H}_1 \to \mathcal{H}_2$, meaning that

$$T_2 = U T_1 U^{-1}.$$

Example 2.19. On $L^2(\mathbb{R}^n)$ the *Fourier transform* is defined as an extension of the integral

$$\hat{f}(\xi) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\, d^n x. \tag{2.18}$$

Although the right-hand side makes sense only if f is integrable, Plancherel's theorem (reviewed in Appendix A.3) says that the map $\mathcal{F} : f \mapsto \hat{f}$ extends by continuity to a unitary operator on $L^2(\mathbb{R}^n)$. ♢

For an inner product space $\mathcal{V}$ which is not complete, we can define its abstract *completion* by a standard metric space construction. Let $\widetilde{\mathcal{V}}$ be the set of equivalence classes of Cauchy sequences in $\mathcal{V}$, where $\{u_j\} \sim \{v_j\}$ if

$$\lim_{j\to\infty} \|u_j - v_j\| = 0.$$

The inner product of two Cauchy sequences is defined in the obvious way by

$$\big\langle \{u_j\}, \{v_j\} \big\rangle := \lim_{j\to\infty} \langle u_j, v_j \rangle. \tag{2.19}$$

One can check that this yields a well-defined inner product, with respect to which $\widetilde{\mathcal{V}}$ is a Hilbert space; see Exercise 2.2. The original space $\mathcal{V}$ is naturally embedded as a dense subspace of $\widetilde{\mathcal{V}}$, by associating a constant sequence to each vector.

Given two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, the *direct sum* is defined as the set of pairs $(u_1, u_2) \in \mathcal{H}_1 \times \mathcal{H}_2$, with

$$\big\langle (u_1, u_2), (v_1, v_2) \big\rangle_{\mathcal{H}_1 \times \mathcal{H}_2} := \langle u_1, v_1 \rangle_{\mathcal{H}_1} + \langle u_2, v_2 \rangle_{\mathcal{H}_2}.$$

To extend this definition to a countable sum requires an extra restriction. For the sequence of Hilbert spaces $\{\mathcal{H}_j\}_{j=1}^{\infty}$, we define

$$\bigoplus_{j=1}^{\infty} \mathcal{H}_j := \Big\{ (u_1, u_2, \dots) : \; u_j \in \mathcal{H}_j, \; \sum \|u_j\|_{\mathcal{H}_j}^2 < \infty \Big\}. \tag{2.20}$$

The assumption on norms guarantees convergence of the inner product defined by

$$\big\langle (u_1, u_2, \dots), (v_1, v_2, \dots) \big\rangle := \sum_{j=1}^{\infty} \langle u_j, v_j \rangle_{\mathcal{H}_j}.$$

We leave to Exercise 2.3 to check that the countable direct sum is complete and therefore a Hilbert space.

2.5 Sobolev Spaces

Many of the applications developed later in this book involve differential operators. One of the primary tools used to analyze such operators is a family of function spaces defined by requiring derivatives up to a certain order to lie in an L^p space. Although Beppo Levi was the first to consider such spaces, in the early twentieth century, they are named for Sergei Sobolev, whose systematic development in 1950 [84] helped to establish their fundamental importance in PDE theory. For simplicity, we will develop only the $p = 2$ case, since this covers all of our applications.

2.5.1 Weak Derivatives

The classical notion of derivative does not apply to L^2 functions, which are not necessarily even defined at all points. To impose the differentiability requirements for Sobolev spaces, we use instead the concept of weak derivatives. These are essentially functions which behave like derivatives in terms of integration by parts.

For an open subset $\Omega \subset \mathbb{R}^n$, let $L^1_{\text{loc}}(\Omega)$ denote the space of *locally integrable* functions, i.e., functions $\Omega \to \mathbb{C}$ which are integrable over compact subsets of Ω. As with the other L^p spaces, functions in L^1_{loc} are taken to be equivalent if they are equal almost everywhere.

For a multi-index $\alpha \in (\mathbb{N}_0)^n$, define the differential operator

$$D^\alpha := \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n},$$

of order $|\alpha| := \alpha_1 + \cdots + \alpha_n$. Let $u \in L^1_{\text{loc}}(\Omega)$. If there exists a function $u^{(\alpha)} \in L^1_{\text{loc}}(\Omega)$ such that

$$\int_\Omega u^{(\alpha)} \psi \, dx = (-1)^{|\alpha|} \int_\Omega u D^\alpha \psi \, d^n x \tag{2.21}$$

for all $\psi \in C_0^\infty(\Omega)$, then we say that u admits a *weak derivative* $D^\alpha u := u^{(\alpha)}$. (The subscript on C_0^∞ indicates compact support.) This is just the standard integration by parts formula, so the weak definition includes classical derivatives as a special case.

Example 2.20. In dimension one, by the Lebesgue differentiation theory, a function which is absolutely continuous is differentiable almost everywhere. Moreover, its derivative is integrable, and the fundamental theorem of calculus relates the derivative to the integral in the usual way. Therefore, if f is absolutely continuous on some interval $I \subset \mathbb{R}$, then the function f' defined in the Lebesgue sense qualifies as a weak derivative.

To see how weak differentiability can fail, consider a function created by splicing two separate functions together,

$$f(x) := \begin{cases} f_+(x), & x > 0, \\ f_-(x), & x < 0, \end{cases}$$

where $f_\pm$ are C^1. For $\psi \in C_0^\infty(\mathbb{R})$, integration by parts gives

$$\begin{aligned} -\int_{-\infty}^{\infty} f \psi' \, dx &= -\int_{-\infty}^{0} f_- \psi' \, dx - \int_0^\infty f_+ \psi' \, dx \\ &= (f_+ - f_-)\psi \Big|_{x=0} + \int_{-\infty}^0 f_-' \psi \, dx + \int_0^\infty f_+' \psi \, dx. \end{aligned}$$

The first term depends on ψ only through $\psi(0)$, and so cannot possibly be recovered from the integral of ψ against another function. Therefore, the weak derivative of f exists only if $f_+(0) = f_-(0)$, and is given in this case by splicing the two derivatives,

$$f'(x) = \begin{cases} f'_+(x), & x > 0, \\ f'_-(x), & x < 0. \end{cases}$$

◊

Example 2.21. For $f \in L^2(\mathbb{R}^n)$, a formal integration by parts in (2.18) gives the formula $\mathcal{F}(D^\alpha f) = (i\xi)^\alpha \hat{f}$, which we can use to construct a weak derivative. Assuming that $\xi^\alpha \hat{f} \in L^2(\mathbb{R}^n)$, set

$$f^{(\alpha)} := \mathcal{F}^{-1}((i\xi)^\alpha \hat{f}).$$

By Plancherel's theorem, for $\psi \in C_0^\infty(\mathbb{R}^n)$,

$$\begin{aligned} \langle f^{(\alpha)}, \psi \rangle &= \langle (i\xi)^\alpha \hat{f}, \hat{\psi} \rangle \\ &= (-1)^{|\alpha|} \langle \hat{f}, (i\xi)^\alpha \hat{\psi} \rangle \\ &= (-1)^{|\alpha|} \langle \hat{f}, \widehat{D^\alpha \psi} \rangle \\ &= (-1)^{|\alpha|} \langle f, D^\alpha \psi \rangle. \end{aligned} \tag{2.22}$$

This shows that $D^\alpha f = f^{(\alpha)}$ as a weak derivative.

This argument can be reversed to prove the converse statement. If $f \in L^2(\mathbb{R}^n)$ has a weak derivative $D^\alpha f$ which is also contained in $L^2(\mathbb{R}^n)$, then $\xi^\alpha f \in L^2(\mathbb{R}^n)$ and $D^\alpha f$ is the inverse Fourier transform of $(i\xi)^\alpha \hat{f}$. ◊

Because $C_0^\infty(\Omega)$ is a dense subset of $L^1(\Omega)$, a weak derivative $D^\alpha u$ defined by (2.21) is unique as an element of $L^1_{\text{loc}}(\Omega)$. It also follows easily from (2.21) that weak differentiation is linear, by the linearity of the integrals.

If a function $u \in L^1_{\text{loc}}(\Omega)$ admits weak derivatives up to order $|\alpha|$, then for $f \in C^\infty(\Omega)$ a simple calculation shows that $D^\alpha(fu)$ exists as a weak derivative and satisfies the standard Leibniz rule,

$$D^\alpha(fu) = \sum_{\beta \le \alpha} \frac{\alpha!}{\beta!(\alpha - \beta)!} (D^\beta f)(D^{\alpha-\beta} u).$$

The fact that we use the same notation for both weak and classical derivatives is justified by the following consistency result.

Lemma 2.22. *A function in $L^1_{\text{loc}}(\Omega)$ is contained in $C^m(\Omega)$ for $m \in \mathbb{N}$ if and only if it is continuous and its weak derivatives to order m exist and are continuous. In this case the weak and classical derivatives coincide.*

Proof If $u \in C^m(\Omega)$, then the classical derivatives satisfy (2.21) by integration by parts. For the converse, it suffices to work on $\mathbb{R}^n$ since C^m is a local condition. Suppose that $u \in C(\mathbb{R}^n)$ has weak derivatives $u^{(\alpha)} \in C(\mathbb{R}^n)$ satisfying (2.21) for $|\alpha| \le m$. Fix $\psi \in C_0^\infty(\mathbb{R}^n)$ with

$$\int_{\mathbb{R}^n} \psi \, d^n x = 1,$$

and set $\psi_\varepsilon(x) := \varepsilon^{-n}\psi(x/\varepsilon)$ for $\varepsilon > 0$. Then $u * \psi_\varepsilon$ is a smooth approximation to u, converging uniformly on compact sets as $\varepsilon \to 0$.

By (2.21),

$$\begin{aligned} u^{(\alpha)} * \psi_\varepsilon &= u * (D_\alpha \psi_\varepsilon) \\ &= D^\alpha (u * \psi_\varepsilon), \end{aligned}$$

where D^α denotes a classical derivative here. Hence, by the assumption that $u^{(\alpha)}$ is continuous, $D^\alpha(u * \psi_\varepsilon)$ converges to $u^{(\alpha)}$ uniformly on compact sets for $|\alpha| \le m$. A standard calculus argument, based on the mean value theorem, shows that for a sequence of C^1 functions, uniform convergence of both functions and derivatives implies that the limit function is also C^1. Applying this inductively in m yields $u \in C^m(\mathbb{R}^n)$. □

2.5.2 H^m Spaces

The *Sobolev spaces* on an open set $\Omega \subset \mathbb{R}^n$ are defined for $m \in \mathbb{N}$ by

$$H^m(\Omega) := \left\{ u \in L^2(\Omega) : \ D^\alpha u \in L^2(\Omega) \text{ for } |\alpha| \le m \right\}, \tag{2.23}$$

where $D^\alpha u$ is defined weakly as in Section 2.5.1. As with L^p spaces, Sobolev functions are identified if they are equal almost everywhere. These spaces are equipped with inner products,

$$\langle u, v \rangle_{H^m} := \sum_{|\alpha| \le m} \langle D^\alpha u, D^\alpha v \rangle, \tag{2.24}$$

and corresponding norms,

$$\|u\|_{H^m} := \left(\sum_{|\alpha| \le m} \|D_\alpha u\|^2 \right)^{1/2}. \tag{2.25}$$

In our notation, $\|\cdot\|_2$ is reserved for to the L^2 norm, assuming the domain Ω is clear from context. The corresponding inner product is denoted by $\langle \cdot, \cdot \rangle$ without subscript. The H^m norms and inner products will always be indicated explicitly, as in (2.24) and (2.25).

Example 2.23. From Example 2.21, we can deduce that

$$H^m(\mathbb{R}^n) = \left\{ f \in L^2(\mathbb{R}^n) : |\xi|^m \hat{f} \in L^2(\mathbb{R}^n) \right\}. \tag{2.26}$$

Note that (2.26) makes sense for all $m \geq 0$, not just for integers. ◊

Example 2.24. We claim that

$$H^1(0,1) = \left\{ f \in AC[0,1] : f' \in L^2(0,1) \right\},$$

where AC denotes the space of absolutely continuous functions. The inclusion

$$\left\{ f \in AC[0,1] : f' \in L^2(0,1) \right\} \subset H^1(0,1)$$

follows immediately from the Lebesgue differentiation theory cited in Example 2.20.

To prove the converse statement, suppose that $f \in H^1(0,1)$. The fact that the weak derivative f' is contained in $L^2(0,1)$ implies that f' is integrable on $(0,1)$. We can thus define its antiderivative by

$$g(x) := \int_0^x f'(t)\, dt.$$

Again by the Lebesgue differentiation theorem, g is absolutely continuous on $[0,1]$, g' exists (in the classical sense) almost everywhere, and $g' = f'$ in the L^2 sense. This implies that

$$\int_0^1 (f-g)'\, \psi\, dx = 0,$$

for all $\psi \in C_0^\infty(\Omega)$, i.e., the weak derivative of $f - g$ is 0. By Lemma 2.22, $f - g$ is classically differentiable and hence constant. This proves that f is absolutely continuous on $[0,1]$. ◊

The definition (2.23) can be extended by replacing the L^2 norms with L^p. This yields a larger family of Sobolev spaces which are denoted by $W^{m,p}$. The use of the letter H for the case $p = 2$ is meant to reflect the following result:

Theorem 2.25. *$H^m(\Omega)$ is a Hilbert space for each $m \in \mathbb{N}$.*

Proof Suppose that $\{u_m\}$ is a Cauchy sequence in $H^m(\Omega)$. Then $\{u_m\}$ is Cauchy with respect to the L^2 norm in particular, so we can define a limit function $u := \lim u_m$ in $L^2(\Omega)$. Similarly, for each α with $|\alpha| \leq m$, $\{D^\alpha u_m\}$ is a Cauchy sequence in $L^2(\Omega)$ and we can define

$$u^{(\alpha)} := \lim_{m\to\infty} D^\alpha u_m$$

in $L^2(\Omega)$.

It remains to show that $u^{(\alpha)}$ is the weak derivative of u. By the definition of $D^\alpha u_m$, given $\psi \in C_0^\infty(\Omega)$,

$$\int_\Omega \psi D^\alpha u_m \, dx = (-1)^{|\alpha|} \int_\Omega u_m D^\alpha \psi \, d^n x.$$

The L^2 convergence allows us to take $m \to \infty$ on both sides to obtain

$$\int_\Omega \psi u^{(\alpha)} \, dx = (-1)^{|\alpha|} \int_\Omega u D^\alpha \psi \, d^n x.$$

Hence the weak derivative $D^\alpha u$ exists for $|\alpha| \le m$ and is given by $u^{(\alpha)} \in L^2(\Omega)$. □

Sobolev spaces are used to measure regularity of functions, providing a more flexible alternative to the family of C^m spaces. We saw in Example 2.24 that H^1 functions are continuous in dimension one. It turns out that Sobolev regularity can be translated back to classical regularity more generally, but the relationship depends on the dimension.

Theorem 2.26 (Sobolev Embedding). *For $\Omega \subset \mathbb{R}^n$, if $m > k + n/2$ then a function in $H^m(\Omega)$ admits a representative in $C^k(\Omega)$.*

Proof Since differentiability is a local property, it suffices to prove this for $\Omega = \mathbb{R}^n$. For $u \in H^m(\mathbb{R}^n)$, let us denote the weak derivatives for $|\alpha| \le m$ by

$$u^{(\alpha)} := \mathcal{F}^{-1}((i\xi)^\alpha \hat{u}).$$

For this proof we will reserve the notation D^α for the classical derivative.

For $|\alpha| \le k$ we have

$$\int_{\mathbb{R}^n} \left|\xi^\alpha \hat{u}\right| d^n\xi \le \int_{\mathbb{R}^n} (1 + |\xi|^2)^{m/2} \left|\hat{u}(\xi)\right| \frac{|\xi|^k}{(1 + |\xi|^2)^{m/2}} \, d^n\xi.$$

The function $|\xi|^k (1 + |\xi|^2)^{-m/2}$ is in L^2 for $m > k + n/2$, so the Cauchy–Schwarz inequality gives

$$\left\|\xi^\alpha \hat{u}\right\|_1 \le C \|u\|_{H^m}.$$

By Riemann–Lebesgue (Lemma A.16), the fact that $\xi^\alpha \hat{u} \in L^1$ implies that $u^{(\alpha)}$ is continuous. Therefore $u^{(\alpha)} = D^\alpha u$ in the classical sense, by Lemma 2.22. □

If a function in H^m admits a continuous representative, then we assume that this representative is chosen by default. Under this convention, the conclusion of the embedding theorem could be restated as

$$H^m(\Omega) \subset C^k(\Omega).$$

Note that Theorem 2.26 covers only differentiability in the interior. The result can be improved to $C^k(\overline{\Omega})$, provided $\partial\Omega$ is sufficiently regular. (See, for example, Evans [29, §5.6.3].)

2.6 Orthogonality

One of the principal reasons why Hilbert spaces are easier to work with than Banach spaces is the concept of orthogonality associated with the inner product. This implies some important results with a geometric flavor analogous to the Euclidean theory of vectors in $\mathbb{R}^n$.

In a Hilbert space $\mathcal{H}$, the *orthogonal complement* of a subset E is defined by

$$E^\perp := \left\{u \in \mathcal{H} : \langle u, v\rangle = 0 \text{ for all } v \in E\right\}.$$

The linearity of the inner product implies that $E^\perp$ is a subspace, even if E is not. Furthermore, since the inner product is easily seen to be continuous by the Cauchy–Schwarz inequality, the space $E^\perp$ is closed. It is useful to note that

$$(E^\perp)^\perp = \overline{\operatorname{span}(E)}, \tag{2.27}$$

where *span* denotes the set of all finite linear combinations of elements; see Exercise 2.4.

As in linear algebra, orthogonal complements provide a tool for decomposing Hilbert spaces as a direct sum of subspaces. The one key difference is that subspaces are automatically closed in finite dimensions, while for infinite-dimensional spaces this is an extra condition.

Theorem 2.27 (Orthogonal Decomposition). *If W is a closed subspace of $\mathcal{H}$, then*

$$\mathcal{H} = W \oplus W^\perp.$$

Proof Since $W \cap W^\perp = \{0\}$, by the positive definiteness of the inner product, it suffices to show that $\mathcal{H} = W + W^\perp$. The strategy comes from the fact that, in the Euclidean case, the projection of $u \in \mathcal{H}$ into W is the vector $w \in W$ which is closest to u, as illustrated in Figure 2.1.

To construct w, let

$$d := \inf_{y\in\mathcal{W}} \|y - u\|,$$

and choose a sequence $\{w_n\}$ in W such that

$$\lim_{n\to\infty} \|w_n - u\| = d. \tag{2.28}$$

We claim that $\{w_n\}$ is Cauchy. To see this, first note that by the parallelogram law,

$$\begin{aligned}\|w_n - w_m\|^2 &= \|(w_n - u) - (w_m - u)\|^2 \\ &= 2\|w_n - u\|^2 + 2\|w_m - u\|^2 - \|w_n + w_m - 2u\|^2.\end{aligned}$$

Since $\frac{1}{2}(w_n + w_m) \in W$,

$$\left\| \tfrac{1}{2}(w_n + w_m) - u \right\| \geq d,$$

and thus

$$\|w_n - w_m\|^2 \leq 2\|w_n - u\|^2 + 2\|w_m - u\|^2 - 4d^2.$$

By (2.28), this shows that

$$\lim_{n,m\to\infty} \|w_n - w_m\| = 0.$$

Hence $\{w_n\}$ is Cauchy and therefore there exists a limit vector $w := \lim w_n$. By the assumption that W is closed, $w \in W$.

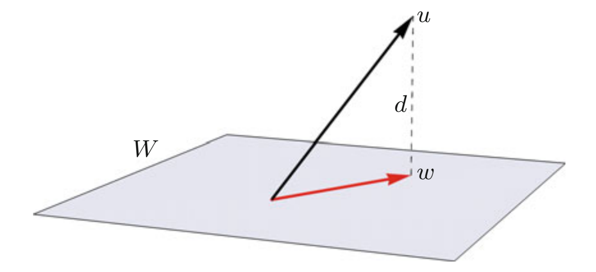


Fig. 2.1 Orthogonal projection into W

To complete the proof, we must show that $v := u - w$ lies in $W^\perp$. Note that $\|v\| = d$ by the construction of w. Let $y \in W$ and $\lambda \in \mathbb{C}$. Because $w - \lambda y \in W$,

$$\begin{aligned}d^2 &\leq \|u - (w - \lambda y)\|^2 \\ &= \|v + \lambda y\|^2 \\ &= d^2 + 2\,\mathrm{Re}[\lambda\langle v, y\rangle] + |\lambda|^2\|y\|^2.\end{aligned}$$

Setting $\lambda = re^{i\theta}$ and taking $r \to 0$ gives

$$\mathrm{Re}\Big[e^{i\theta}\langle v, y\rangle\Big] \geq 0.$$

This inequality holds for all θ, implying that $\langle v, y\rangle = 0$. Since y was an arbitrary element of W, it follows that $v \in W^\perp$. □

An *orthogonal projection* on $\mathcal{H}$ is an operator $P \in \mathcal{L}(\mathcal{H})$ such that $P^2 = P$ and $\mathcal{H} = \text{range}(P) \oplus \ker(P)$. Theorem 2.27 could be paraphrased as the statement that for each closed subspace $W \subset \mathcal{H}$ there exists an orthogonal projection with range equal to W.

Orthogonal decomposition plays an important role in the proof of the following result, which identifies elements of the dual space $\mathcal{H}' := \mathcal{L}(\mathcal{H}, \mathbb{C})$ with vectors in $\mathcal{H}$.

Theorem 2.28 (Riesz Lemma). *For each $F \in \mathcal{H}'$, there exists a unique $v \in \mathcal{H}$ such that*

$$F(u) = \langle v, u \rangle$$

for all $u \in \mathcal{H}$. Furthermore, $\|F\| = \|v\|$.

Proof By the linearity and continuity of F, $\ker(F)$ is a closed subspace, so Theorem 2.27 gives a decomposition

$$\mathcal{H} = \ker(F) \oplus \ker(F)^{\perp}.$$

If $\ker(F)^{\perp} = \{0\}$, then $F = 0$ and the result is obtained by setting $v = 0$.

Assume that $\ker(F)^{\perp}$ contains a vector $w \neq 0$. For $u \in \mathcal{H}$, note that $F(w)u - F(u)w \in \ker(F)$. Therefore, since w is orthogonal to $\ker(F)$,

$$\begin{aligned} 0 &= \langle w, F(w)u - F(u)w \rangle \\ &= F(w)\langle w, u \rangle - F(u)\|w\|^2. \end{aligned}$$

Since $w \neq 0$ we can then define

$$v := \frac{\overline{F(w)}}{\|w\|^2} w,$$

which yields the desired relation

$$F(u) = \langle v, u \rangle.$$

The uniqueness of v follows immediately from the positive definiteness of the inner product.

To see that $\|F\| = \|v\|$, observe that $F(v) = \|v\|^2$ implies that $\|F\| \geq \|v\|$. On the other hand, the Cauchy–Schwarz inequality gives

$$|F(u)| \leq \|u\| \|v\|$$

for all u, so $\|F\| \leq \|v\|$. □

Note that the canonical map $\mathcal{H}' \to \mathcal{H}$ is conjugate-linear rather than linear. One immediate application of the Riesz lemma is the following association of linear maps to sesquilinear forms, whose proof is left to Exercise 2.6.

Corollary 2.29. *Let $\eta : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ be a sesquilinear form which is bounded in the sense that*

$$\|\eta\| := \sup_{v,w\neq 0} \frac{|\eta(v,w)|}{\|v\|\|w\|} < \infty.$$

Then there is a uniquely defined operator $T \in \mathcal{L}(\mathcal{H})$ such that

$$\eta(v,w) = \langle v, Tw\rangle.$$

The Riesz lemma also allows us to simplify the definition (2.11) of weak operator convergence in the Hilbert space context. For operators in $\mathcal{L}(\mathcal{H})$, $T_n \to T$ in the weak operator sense if and only if

$$\lim_{n\to\infty} \langle v, T_n w\rangle = \langle v, Tw\rangle$$

for all $v, w \in \mathcal{H}$. The weak operator convergence $S^k \to 0$ in Example 2.11 follows immediately from this formulation.

The identification of elements of $\mathcal{H}$ with functionals via the Riesz lemma also provides a new topology on $\mathcal{H}$, derived from the weak operator topology on $\mathcal{H}'$. For a sequence $\{w_n\} \in \mathcal{H}$, we say that $w_n \to w$ *weakly* if

$$\lim_{n\to\infty} \langle w_n, v\rangle = \langle w, v\rangle$$

for all $v \in \mathcal{H}$.

Example 2.30. In $L^2(\mathbb{R})$ let $f_n := \chi_{[n,n+1]}$. For $u \in L^2(\mathbb{R})$, Cauchy–Schwarz gives the estimate

$$|\langle u, f_n\rangle| \le \|u\chi_{[n,n+1]}\|_2.$$

Since $\sum_{n\in\mathbb{Z}} \|u\chi_{[n,n+1]}\|_2^2 = \|u\|_2^2 < \infty$, this shows that $\langle u, f_n\rangle \to 0$ as $n \to \infty$. Hence, $f_n \to 0$ weakly even though $\|f_n\|_2 = 1$ for all n. ♢

2.7 Orthonormal Bases

A Hilbert space is called *separable* if it admits a countable dense subset. For example, in $L^2(\mathbb{R}^n)$ we can produce a countable dense set consisting of step functions built using only rational numbers. Most Hilbert spaces of practical interest are separable, including all of the cases encountered in this book.

To simplify the exposition, we will assume that $\mathcal{H}$ is infinite-dimensional but separable throughout this section. The same arguments apply in finite dimensions, but there are some technicalities, for example in the use of the term "sequence," which is normally required to be infinite. The non-separable case requires a more substantial revision of the definition of a basis, which we will not go into here.

Definition 2.31. A sequence $\{e_1, e_2, \ldots\} \subset \mathcal{H}$ is *orthonormal* if

$$\langle e_i, e_j \rangle = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

An *orthonormal basis* is an orthonormal sequence whose span is dense in $\mathcal{H}$.

Example 2.32. The Fourier basis for the space $L^2(0, 2\pi)$ is given by $\{\phi_k\}_{k\in\mathbb{Z}}$ where

$$\phi_k(\theta) := \frac{1}{\sqrt{2\pi}} e^{ik\theta}. \tag{2.29}$$

It is easy to check that the sequence $\{\phi_k\}$ is orthonormal. To see that it forms a basis, it is convenient to replace the interval $(0, 2\pi)$ with the quotient $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. Functions on $\mathbb{T}$ are interpreted as periodic functions on $\mathbb{R}$, so the spaces $L^2(\mathbb{T})$ and $L^2(0, 2\pi)$ are naturally isomorphic.

A standard argument using the Dirichlet kernel shows that the Fourier series of a function in $C^1(\mathbb{T})$ converges uniformly, and hence in L^2. Since $C^1(\mathbb{T})$ is dense in $L^2(\mathbb{T})$, this implies that $\{\phi_k\}$ is an orthonormal basis.

A similar argument applies to the n-dimensional torus $\mathbb{T}^n := \mathbb{R}^n/(2\pi\mathbb{Z})^n$. The Fourier orthonormal basis for $L^2(\mathbb{T}^n)$ is given by $\{\phi_k\}_{k\in\mathbb{Z}^n}$, with

$$\phi_k(\theta) := (2\pi)^{-n/2} e^{ik\cdot\theta}.$$

By restricting and changing coordinates as needed, we can derive from this construction a Fourier basis for the L^2 space of a bounded rectangle in $\mathbb{R}^n$. ♢

Clearly, a Hilbert space that admits an orthonormal basis is separable, because the set of linear combinations of basis elements with rational coefficients furnishes a countable dense subset. The converse also holds, as seen in the following:

Theorem 2.33. *A separable Hilbert space admits an orthonormal basis (in the sense defined above).*

Proof Assume that $\{f_j\}$ is a countable dense subset of $\mathcal{H}$. For each $n \in \mathbb{N}$, an application of the Gram–Schmidt process from linear algebra to the set $\{f_1, \ldots, f_n\}$ produces an orthonormal set $\{e_1, \ldots, e_{m_n}\}$ such that

$$\operatorname{span}\{f_1, \ldots, f_n\} = \operatorname{span}\{e_1, \ldots, e_{m_n}\}.$$

The density of $\{f_j\}$ in $\mathcal{H}$ then implies the density of $\operatorname{span}\{e_j\}$. □

Given an orthonormal sequence $\{e_j\}_{j=1}^{\infty}$, let P_n denote the orthogonal projection onto $\mathrm{span}\{e_1, \ldots, e_n\}$, for each $n \in \mathbb{N}$. It is easy to check, by the construction in Theorem 2.27, that

$$P_n v = \sum_{j=1}^{n} \langle e_j, v\rangle e_j. \tag{2.30}$$

To decide if $\{e_j\}$ forms a basis, we need to investigate the convergence of $P_n v$ as $n \to \infty$.

Theorem 2.34 (Bessel's Inequality). *Suppose $\{e_j\}_{j=1}^{\infty}$ is an orthonormal sequence in $\mathcal{H}$. For $v \in \mathcal{H}$,*

$$\sum_{j=1}^{\infty} \left|\langle e_j, v\rangle\right|^2 \le \|v\|^2,$$

with equality if and only if $P_n v \to v$ as $n \to \infty$.

Proof Since P_n is an orthogonal projection,

$$\begin{aligned}\|v\|^2 &= \left\|v - P_n v\right\|^2 + \left\|P_n v\right\|^2 \\ &= \left\|v - P_n v\right\|^2 + \sum_{j=1}^{n} \left|\langle e_j, v\rangle\right|^2.\end{aligned}$$

The sum converges as $n \to \infty$, because its terms are all positive. Hence,

$$\|v\|^2 = \lim_{n\to\infty} \left\|v - P_n v\right\|^2 + \sum_{j=1}^{\infty} \left|\langle e_j, v\rangle\right|^2,$$

from which both claims follow directly. □

One immediate application of Bessel's inequality is the following set of characterizations of a basis.

Theorem 2.35 (Basis Criteria). *For an orthonormal sequence $\{e_1, e_2, \ldots\}$ in a separable Hilbert space $\mathcal{H}$, the following statements are equivalent:*

(a) *The sequence is an orthonormal basis.*
(b) *The only vector perpendicular to e_j for all j is zero.*
(c) *$P_n \to I$ in the strong operator sense as $n \to \infty$.*

Proof To prove that (a) implies (b), note that

$$\{e_1, \ldots\}^{\perp} = (\mathrm{span}\{e_1, \ldots\})^{\perp}.$$

Thus, for a basis $\{e_1, \ldots\}^{\perp} = \{0\}$.

Assume (b) holds. For $v \in \mathcal{H}$, the sequence $\{P_n v\}$ is Cauchy by Bessel's inequality, since

$$\left\| P_n v - P_m v \right\|^2 = \sum_{j=n}^{m} \left| \langle e_j, v \rangle \right|^2$$

for $n \leq m$. Therefore $P_n v$ converges to some limit $\tilde{v}$. For $j \leq n$,

$$\left\langle e_j, v - P_n v \right\rangle = 0.$$

Taking $n \to \infty$ then yields

$$\langle e_j, v - \tilde{v} \rangle = 0,$$

so that $v - \tilde{v}$ is orthogonal to all of the e_j. Hence $v = \tilde{v}$ by (b), and thus $P_n v \to v$.

For the final claim that (c) implies (a), we simply note that $P_n v \in \operatorname{span}\{e_j\}$ by (2.30). □

From the combination of Theorems 2.34 and 2.35, we immediately derive the following result on basis expansion.

Corollary 2.36 (Parseval's Theorem). *If $\{e_j\}$ is an orthonormal basis for $\mathcal{H}$, then each $v \in \mathcal{H}$ can be represented as a convergent series*

$$v = \sum_{j=1}^{\infty} \langle e_j, v \rangle e_j.$$

Furthermore,

$$\|v\|^2 = \sum_{j=1}^{\infty} \left| \langle e_j, v \rangle \right|^2.$$

2.7.1 Weak Sequential Compactness

A subset $K \subset \mathcal{H}$ is called *sequentially compact* if every infinite sequence in K admits a subsequence converging in K. (Because the topology is metric, this is equivalent to the usual topological definition of compactness.) Under the assumption that $\mathcal{H}$ is infinite-dimensional, a closed and bounded subset is not necessarily compact. For example, an orthonormal basis is contained in the closed unit ball and admits no convergent subsequence.

This picture changes, however, if we consider weak convergence. The following result could be paraphrased as the statement that a closed and bounded subset of $\mathcal{H}$ is "weakly sequentially compact." We will exploit the existence result from Theorem 2.33 to produce a rather simple proof in the case of a separable Hilbert space.

Theorem 2.37 (Alaoglu's Theorem). *Suppose $\{u_k\}$ is a sequence in a separable Hilbert space $\mathcal{H}$ satisfying $\|u_k\| \le M$ for all k. Then there exists a subsequence of $\{u_k\}$ that converges weakly to some $w \in \mathcal{H}$ with $\|w\| \le M$.*

Proof Assume that u_k satisfies $\|u_k\| \le M$ for each k. Let $\{e_j\}$ be an orthonormal basis for $\mathcal{H}$. By assumption,

$$|\langle e_1, u_k\rangle| \le M.$$

By the Bolzano–Weierstrass theorem, there exists a subsequence $\{u_{1,k}\}$ such that $\langle e_1, u_{1,k}\rangle$ converges to some $a_1 \in \mathbb{C}$ as $k \to \infty$. We can then find a further subsequence $\langle e_2, u_{2,k}\rangle$ converges to a_2, and so on. Setting $w_k = u_{k,k}$ gives a subsequence for which

$$\lim_{k\to\infty} \langle e_j, w_k\rangle = a_j, \tag{2.31}$$

for each $j \in \mathbb{N}$.

The next step is to estimate the coefficients a_j to show that they correspond to a vector in $\mathcal{H}$. By the uniform bound on $\|u_k\|$,

$$\sum_{j=1}^{n} \left|\langle e_j, w_k\rangle\right|^2 \le M^2,$$

for all k and n. Taking $k \to \infty$ and then $n \to \infty$ gives

$$\sum_{j=1}^{\infty} |a_j|^2 \le M^2.$$

This implies that $w := \sum_j a_j e_j$ is well defined in $\mathcal{H}$ and satisfies $\|w\| \le M$. By (2.31),

$$\lim_{k\to\infty} \langle e_j, w_k\rangle = \langle e_j, w\rangle$$

for all j. This is sufficient to guarantee weak convergence $w_k \to w$ (see Exercise 2.8). □

2.8 Exercises

2.1. For normed vector spaces $\mathcal{V}$ and $\mathcal{W}$, prove that a linear map $T : \mathcal{V} \to W$ is bounded if and only if it is continuous.

2.2. For the Hilbert space completion defined in Section 2.4, show that (2.19) defines an inner product and that this makes $\widetilde{\mathcal{V}}$ a Hilbert space.

2.3. Show that the countable direct sum defined in (2.20) is a Hilbert space.

2.4. Prove the claim (2.27): For a Hilbert space $\mathcal{H}$ and a subset $E \subset \mathcal{H}$,

$$(E^\perp)^\perp = \overline{\operatorname{span}(E)}.$$

2.5. For $T \in \mathcal{L}(\mathcal{H})$, prove that

$$\|T\| = \sup_{v,w\neq 0} \frac{|\langle v, Tw\rangle|}{\|v\|\|w\|}.$$

2.6. Prove Corollary 2.29: Given a bounded sesquilinear form $\eta : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$, there is a unique bounded operator T such that

$$\eta(v, w) = \langle v, Tw\rangle$$

for all $v, w \in \mathcal{H}$.

2.7. In a Hilbert space $\mathcal{H}$, prove that an orthonormal sequence $\{e_n\}$ converges weakly to 0.

2.8. Let $\{e_j\}$ be an orthonormal basis in a Hilbert space $\mathcal{H}$. If $w \in \mathcal{H}$ and $\{w_n\}$ is a bounded sequence such that

$$\lim_{n\to\infty} \langle w_n, e_j\rangle = \langle w, e_j\rangle$$

for each j, prove that $w_n \to w$ weakly.

2.9. Suppose that $A : \mathcal{H} \to \mathcal{H}$ is a linear map satisfying

$$\langle Av, w\rangle = \langle v, Aw\rangle$$

for all $v, w \in \mathcal{H}$. Use the uniform boundedness principle to prove that A is bounded.

2.10. Suppose that $P \in \mathcal{L}(\mathcal{H})$ satisfies $P^2 = P$ and

$$\langle Pv, w\rangle = \langle v, Pw\rangle$$

for all $v, w \in \mathcal{H}$. Prove that P is an orthogonal projection (as defined in Section 2.6) onto the range of P.

2.11. The *Schur test* gives a useful criterion for establishing the boundedness of operators defined by integral kernels. Suppose that $K(\cdot, \cdot)$ is a measurable function on $\mathbb{R}^{2n}$, and there exists a constant C such that

$$\int_{\mathbb{R}^n} |K(x, y)| \, d^n y \leq C, \qquad \int_{\mathbb{R}^n} |K(x, y)| \, d^n x \leq C.$$

Prove that the operator on $L^2(\mathbb{R}^n)$ defined by

$$Tf(x) := \int_{\mathbb{R}^n} K(x, y) f(y) \, d^n y$$

is bounded, with $\|T\| \leq C$. [Hint: first use the Cauchy–Schwarz inequality to estimate $|Tf(x)|^2$, then integrate over x and use Fubini's theorem to switch the order of integrations.]

Notes

Metric space topology is covered in most undergraduate analysis texts; a classic source is Rudin [77, Chapter 2]. An introduction to L^p spaces can be found in basic treatments of measure theory, such as Folland [31, Chapter 6] or Royden [76, Chapter 6]. Additional background material on measure theory and integration is provided in Appendix A.1.

For a more complete introduction to topologies on the space of bounded operators, we refer the reader to Reed and Simon [69, Chapter VI]. The simple proof of uniform boundedness (Theorem 2.13), which was adapted from Sokal [85], uses a technique from the earliest versions of the proof by Hahn and Banach in 1922. In 1927, Banach and Steinhaus gave a different version of the proof, relying on the Baire category theorem (see [69, §III.5]). This route is more efficient, in the sense that Baire category has other important corollaries, but less direct.

There are many excellent sources for the basic material on Hilbert spaces from Sections 2.4, 2.6, and 2.7. See, for example, MacCluer [60, Chapter 1] or Stein and Shakarchi [87, Chapter 4].

While Sobolev spaces are not usually covered in basic functional analysis texts, this is standard background for PDE theory. See Borthwick [13, Chapter 10] for a gentle introduction, and Evans [29, Chapter 5] or Gilbarg and Trudinger [36, Chapter 7] for a more comprehensive treatment. We will develop more aspects of this theory in Chapters 6 and 9.

Chapter 3
Operators



A fundamental issue in the spectral theory of differential operators is the fact that an L^2 function space is the natural Hilbert space for many applications, and yet L^2 functions are not differentiable. Before developing the spectral theory, we must deal with the basic problem of what it means for a differential operator to act on an L^2 space.

One way to circumvent this issue is to consider the inverse of the operator. In many classical PDE problems, differential operators have integral solution kernels called Green's functions. These solution kernels, introduced in the 1830s by George Green, define operators which are generally bounded in the L^2 sense. Early spectral theory, including David Hilbert's foundational work in the early 1900s, focused on the case of integral operators, and thus avoided the technical difficulties associated with differential operators.

For the development of quantum mechanics, a more direct understanding of the spectral theory of the differential operator itself was required. John von Neumann and Marshall Stone addressed this challenge in the late 1920s, by developing the notion of an "unbounded" operator. The innovation they proposed was to relax the assumption that the domain of the operator is the full space. Instead, each operator carries its own domain subspace. This causes some unavoidable technical complications, which we will discuss in this chapter.

3.1 Unbounded Operators

The operator we will be most concerned with later in the book is the *Laplacian*, named for Pierre-Simon Laplace. This is the second order differential operator on $\mathbb{R}^n$ given by

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1_3

$$\Delta := \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}. \tag{3.1}$$

Let $\Omega \subset \mathbb{R}^n$ be an open subset (with Lebesgue measure implied by default). To define Δ as an operator on $L^2(\Omega)$, where Ω is an open subset of $\mathbb{R}^n$, we need to assign it a domain. We could choose $C_0^\infty(\Omega)$, for example, or a subspace of $C^\infty(\overline{\Omega})$ defined by imposing some boundary condition. These choices will result in operators with different spectral properties.

To proceed, we need to refine our usage of the term "operator." Although this term generally refers to a linear map between vector spaces, as in Section 2.3, we will adopt the following more specialized definition in the context of Hilbert spaces.

Definition 3.1. An *operator* T on a Hilbert space $\mathcal{H}$ is a linear map $T : \mathcal{D}(T) \mapsto \mathcal{H}$, where $\mathcal{D}(T)$ is a dense subspace of $\mathcal{H}$, called the *domain* of T.

Making the domain specification implicit to the operator helps keep the notation cleaner, but there is a potential for confusion in the case of differential operators. For example, although the action of Δ on $C^\infty(\Omega)$ is unambiguously defined by (3.1), the Laplacian has many different realizations as an operator on $L^2(\Omega)$. The choice of domain will need either to be clear from context or indicated with some alternate notation.

The density requirement on the domain is a matter of convention. We choose to include this condition in the definition because most of the results we discuss will require it as a hypothesis.

In order to simplify the presentation, we focus on operators on a single Hilbert space $\mathcal{H}$ here, but Definition 3.1 could obviously be applied to maps from one Hilbert space to another. The generalization of the basic results of this section to the case $T : \mathcal{H}_1 \to \mathcal{H}_2$ is straightforward. It simply requires us to distinguish between the two different norms and inner products.

The definition of a bounded operator from Section 2.3 extends to the context of Definition 3.1: an operator T is *bounded* if

$$\sup_{v \in \mathcal{D}(T) \setminus \{0\}} \frac{\|Tv\|}{\|v\|} < \infty,$$

and *unbounded* otherwise. Note that a bounded operator admits a unique continuous extension to the full space $\mathcal{H}$, since $\mathcal{D}(T)$ is dense. When T is bounded we will assume that $\mathcal{D}(T) = \mathcal{H}$ by default.

The notation $\mathcal{L}(\mathcal{H})$ introduced in Section 2.3 is reserved for the algebra of bounded operators. There is no corresponding notation for unbounded operators, as they do not form an algebra or even a vector space. Adding or composing unbounded operators is possible only if the domains line up appropriately. We can always add a bounded perturbation to an unbounded operator without changing the original domain, but an unbounded perturbation requires more care. We will discuss various examples in Chapter 7.

Example 3.2. Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space. For a measurable function $f : X \to \mathbb{C}$, we define the (possibly unbounded) multiplication operator on $L^2(X, d\mu)$,

$$M_f : v \mapsto fv, \tag{3.2}$$

with the domain,

$$\mathcal{D}(M_f) := \{v \in L^2(X, d\mu) : fv \in L^2(X, d\mu)\}.$$

We claim that this domain is dense; see Exercise 3.12. The argument from Example 2.8 shows that M_f is bounded if and only if $f \in L^\infty(X, d\mu)$, with

$$\|M_f\| = \|f\|_\infty.$$

$\diamondsuit$

Example 3.3. The Fourier transform introduced in Example 2.19 allows us to conjugate the Laplacian (3.1) to a multiplication operator. For $\psi \in C_0^\infty(\mathbb{R}^n)$, integration by parts shows that

$$\mathcal{F}(-\Delta\psi) = |\xi|^2\hat{\psi}(\xi).$$

Since the $\mathcal{F}$ is unitary as a map $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$, the operator $-\Delta$ is conjugate to the multiplication operator $M_{|\xi|^2}$. (It is traditional to include a minus sign with Δ in many applications, and the operator $-\Delta$ is also commonly referred to as the Laplacian.)

If the domain of $M_{|\xi|^2}$ is defined as in Example 3.2, then the corresponding domain for $-\Delta$ is

$$\mathcal{D}(-\Delta) = \left\{f \in L^2(\mathbb{R}^n) : |\xi|^2\hat{f} \in L^2(\mathbb{R}^n)\right\}.$$

This is the Sobolev space $H^2(\mathbb{R}^n)$ defined in Section 2.5. $\diamondsuit$

3.2 Adjoints

Let us first limit our attention to the bounded case. Given $v \in \mathcal{H}$ and $T \in \mathcal{L}(\mathcal{H})$, we can define a bounded functional $u \mapsto \langle v, Tu\rangle$. The Riesz lemma (Theorem 2.28) then yields a unique vector w such that

$$\langle v, Tu\rangle = \langle w, u\rangle,$$

for $u \in \mathcal{H}$. The *adjoint* of T is the map $T^* : v \mapsto w$, which is easily seen to be linear. In other words, T^* is the unique linear operator such that

$$\langle v, Tu \rangle = \langle T^* v, u \rangle, \tag{3.3}$$

for all $u \in \mathcal{H}$. From the formula for the operator norm developed in Exercise 2.5, we can see immediately that T^* is bounded, with

$$\|T^*\| = \|T\|, \qquad \|T^*T\| = \|T\|^2. \tag{3.4}$$

Recall from Section 2.4 that a map $F \in \mathcal{L}(\mathcal{H})$ is an isometry if and only if it preserves the inner product, i.e.,

$$\langle Fu, Fv \rangle = \langle u, v \rangle.$$

By the definition of the adjoint, the isometry condition is equivalent to

$$F^* F = I. \tag{3.5}$$

If $U \in \mathcal{L}(\mathcal{H})$ is unitary, i.e., a bijective isometry, then the condition (3.5) implies that $U^{-1} = U^*$. The converse statement, that $U^{-1} = U^*$ implies U is unitary, is also clear. Therefore, unitarity is equivalent to the conditions

$$U^* U = I \quad \text{and} \quad UU^* = I. \tag{3.6}$$

3.2.1 Adjoints of Unbounded Operators

In the unbounded case, we still base the definition of the adjoint on the formula (3.3), but the issue of domains adds a new consideration. In order to apply the Riesz lemma, we need the functional $\langle v, T(\cdot) \rangle$ to be bounded on $\mathcal{D}(T)$. This may not be true for all vectors v, which puts a constraint on the domain of T^*.

When $\langle v, T(\cdot) \rangle$ is a bounded functional, the construction of $T^* v$ works just as above. Because $\mathcal{D}(T)$ is dense, the functional extends by continuity to all $\mathcal{H}$. The Riesz lemma thus yields a unique vector, which we take to be $T^* v$. This leads us to make the following:

Definition 3.4. The *adjoint* of an operator $T : \mathcal{D}(T) \to \mathcal{H}$ is the unique linear map T^* defined by the condition that

$$\langle v, Tu \rangle = \langle T^* v, u \rangle$$

for all $u \in \mathcal{D}(T)$ and $v \in \mathcal{D}(T^*)$, where

$$\mathcal{D}(T^*) := \big\{ v \in \mathcal{H} : \ u \mapsto \langle v, Tu \rangle \text{ is a bounded functional on } \mathcal{D}(T) \big\}. \tag{3.7}$$

One potential issue here is that $\mathcal{D}(T^*)$ need not be dense. In fact, it is possible to have $\mathcal{D}(T^*) = \{0\}$. This technicality will be resolved in Section 3.3, when we will develop a basic criterion for the adjoint to be densely defined, which will be satisfied for all of the operators considered later in the book.

Example 3.5. For the multiplication operator M_f introduced in Example 3.2, we have $\mathcal{D}(M_f^*) = \mathcal{D}(M_f)$ and

$$M_f^* = M_{\overline{f}}.$$

The proof is included in Exercise 3.12. ◊

Example 3.6. On $L^2[0,1]$ consider the operator $T := d/dx$ with domain $\mathcal{D}(T) = C^1[0,1]$. For $u, v \in \mathcal{D}(T)$, integration by parts gives

$$\langle u, Tv\rangle = -\langle Tu, v\rangle + \overline{u(1)}v(1) - \overline{u(0)}v(0).$$

The pairing $v \mapsto \langle Tu, v\rangle$ clearly extends to a continuous functional on $L^2(0,1)$, whereas the evaluation map $v \mapsto v(x_0)$ does not. The boundedness of the functional $v \mapsto \langle u, Tv\rangle$, which is required for $u \in \mathcal{D}(T^*)$, thus implies the boundary conditions $u(0) = u(1) = 0$.

We claim that the full adjoint domain is

$$\mathcal{D}(T^*) = \left\{ f \in AC[0,1] : \; f(0) = f(1) = 0, \; f' \in L^2[0,1] \right\}, \tag{3.8}$$

where $AC[0,1]$ denotes the space of absolutely continuous functions. From the Lebesgue differentiation theory (see, e.g., [87, §3.2, Thm. 3.11]), we recall that for $f \in AC[0,1]$, we have that f is differentiable a.e., $f' \in L^1[0,1]$, and

$$f(x) = f(0) + \int_0^x f'\, dt \tag{3.9}$$

for all $x \in [0,1]$. Conversely, for $g \in L^1(0,1)$ the function defined by

$$f(x) := \int_0^x g\, dt$$

is absolutely continuous, with $f' = g$ a.e.

Suppose that a function $f \in AC[0,1]$ satisfies $f(0) = f(1) = 0$ and $f' \in L^2[0,1]$. For $\phi \in C^1[0,1]$, the product ϕf is also absolutely continuous, so (3.9) implies that

$$\int_0^1 (\phi f)'\, dx = 0,$$

by the endpoint condition on f. Therefore,

$$\langle f, T\phi\rangle = \langle -f', \phi\rangle$$

for all $\phi \in \mathcal{D}(T)$. By the assumption that $f' \in L^2(0,1)$, this shows that $f \in \mathcal{D}(T^*)$ and

$$T^* f = -f'.$$

To prove the converse statement, assume that $f \in \mathcal{D}(T^*)$ and set $g = T^* f$. By the definition of T^*, we have

$$\langle f, T\phi\rangle = \langle g, \phi\rangle. \tag{3.10}$$

for all $\phi \in \mathcal{D}(T)$. The function g is integrable, because $g \in L^2[0,1]$ and $[0,1]$ is compact. We can thus define a function $h \in AC[0,1]$ by

$$h(x) := \int_0^x g(t)\,dt,$$

satisfying $h' = g$ a.e. Applying (3.9) to the product $\bar{h}\phi$ gives

$$\int_0^1 (\bar{h}\phi)'\,dt = \overline{h(1)}\phi(1),$$

and therefore,

$$\langle g, \phi\rangle + \langle h, \phi'\rangle = \overline{h(1)}\phi(1).$$

By (3.10), it follows that

$$\langle f + h, T\phi\rangle = \overline{h(1)}\phi(1) \tag{3.11}$$

for all $\phi \in \mathcal{D}(T)$.

Setting $\phi \equiv 1$ in (3.11) shows that $h(1) = 0$, and thus

$$f + h \in \text{range}(T)^\perp.$$

Since $\text{range}(T) = C[0,1]$, which is dense in $L^2(0,1)$, $\text{range}(T)^\perp = \{0\}$. Therefore $f = -h$, implying that f is absolutely continuous, $f' \in L^2(0,1)$ and $f(0) = f(1) = 0$.

This completes the justification of (3.8). We will consider the calculation of $\mathcal{D}(T^*)$ for some other choices of $\mathcal{D}(T)$ in Exercise 3.1. ♢

We conclude this section by noting a basic relationship between the range of an operator and the kernel of its adjoint.

Lemma 3.7. *For an operator T on a Hilbert space $\mathcal{H}$,*

$$\ker(T^*) = \operatorname{range}(T)^{\perp}.$$

This is a standard fact in linear algebra, and the argument is essentially the same as in finite dimensions. The proof is left to Exercise 3.3.

3.3 Closed Operators

Domains of operators are often difficult to specify precisely, and we frequently need to consider expanding the given domain of an operator. An operator S is said to be an *extension (of an operator)* of T if

$$\mathcal{D}(T) \subset \mathcal{D}(S) \text{ and } S|_{\mathcal{D}(T)} = T.$$

The standard shorthand for this relationship is

$$T \subset S.$$

Note that it follows immediately from the definition of the adjoint that if $T \subset S$, then $S^* \subset T^*$.

Because unbounded operators are discontinuous, extensions are not guaranteed to exist, nor are they necessarily unique. To clarify the situation, it is helpful to consider the *graph* of the operator,

$$\Gamma(T) := \{(u, Tu) : \ u \in \mathcal{D}(T)\} \subset \mathcal{H} \times \mathcal{H}.$$

Using the graph, we can introduce a notion which sheds light on both the extension problem and the issue of densely defined adjoints from Section 3.2.

Definition 3.8. An operator is *closed* if its graph is closed as a subspace of $\mathcal{H} \times \mathcal{H}$.

Note that the closed graph condition does not require $\mathcal{D}(T)$ to be closed; this would be the case only if $\mathcal{D}(T) = \mathcal{H}$. The closure of an operator should be thought of as a weak form of continuity. The statement that "T is closed" means precisely that if $\{u_n\} \subset \mathcal{D}(T)$ and both sequences $\{u_n\}$ and $\{Tu_n\}$ are convergent in $\mathcal{H}$, then $\lim u_n \in \mathcal{D}(T)$ and

$$T\Big(\lim_{n\to\infty} u_n\Big) = \lim_{n\to\infty} Tu_n.$$

The difference between this and continuity lies in the fact that convergence of $\{Tu_n\}$ is part of the hypothesis rather than part of the conclusion. Bounded operators are therefore automatically closed.

Example 3.9. Let $T = d/dx$ acting on $L^2(-1, 1)$ with domain $\mathcal{D}(T) = C^1[-1, 1]$. Although this might seem to be a reasonable choice of domain, T is not closed. For example, set $f_\varepsilon(x) := \sqrt{x^2 + \varepsilon}$ for $\varepsilon > 0$. As $\varepsilon \to 0$,

$$f_\varepsilon(x) \to |x|, \quad f_\varepsilon'(x) \to \operatorname{sign}(x),$$

with convergence in $L^2(-1, 1)$ in both cases. Since $|x| \notin \mathcal{D}(T)$, the graph of T is not closed. ♢

Example 3.10. Consider the multiplication operator M_f introduced in Example 3.2, with $\mathcal{D}(M_f) = \{v \in L^2(X, d\mu) : fv \in L^2(X, d\mu)\}$. To prove closure, consider a sequence $\{u_n\} \subset \mathcal{D}(M_f)$ such that

$$u_n \to u \in L^2(X, d\mu) \quad \text{and} \quad fu_n \to w \in L^2(X, d\mu).$$

By a standard measure theory argument, a convergent sequence in L^2 has a subsequence that converges pointwise almost everywhere. Thus, by passing to subsequence, we can assume that the convergence $u_n \to u$ and $fu_n \to w$ holds pointwise almost everywhere. It follows that $w = fu$ almost everywhere, which implies that $u \in \mathcal{D}(M_f)$ and $M_f u = w$. The operator M_f is therefore closed on this domain.

From the conjugation of the Laplacian on $\mathbb{R}^n$ to a multiplication operator described in Example 3.3, we can also deduce that $-\Delta$ is closed on the domain $\mathcal{D}(-\Delta) = H^2(\mathbb{R})$. ♢

3.3.1 *Closable Operators*

One obvious way to try to extend an operator that is not closed is to consider the closure of its graph. This technique does not always work, however, because the closure of the graph is not necessarily the graph of an operator. It might, for example, contain a pair $(0, v)$ with $v \neq 0$.

We say that T is *closable* if $\overline{\Gamma(T)}$ is the graph of an operator, and define in this case the *closure* $\overline{T}$ by

$$\Gamma(\overline{T}) = \overline{\Gamma(T)}.$$

Obviously $T \subset \overline{T}$, so a closable operator admits at least one closed extension. The converse is also true, as the following result shows.

Lemma 3.11. *If an operator T admits a closed extension S, then T is closable and $\overline{T} \subset S$.*

Proof Suppose T admits a closed extension S. Then, because $\overline{\Gamma(T)}$ is the smallest closed set containing $\Gamma(T)$,

$$\overline{\Gamma(T)} \subset \Gamma(S).$$

It is easy to check that this implies that $\overline{\Gamma(T)}$ is the graph of an operator which is the restriction of S to a possibly smaller domain. Hence T is closable and $\overline{T} \subset S$. □

We should remark that the hypothesis of Lemma 3.11, the existence of at least one closed extension, is often taken as the definition of closable. The closure of an operator could also be defined as its smallest closed extension, without actually mentioning the graph.

There is a simple relationship between the graph of an operator and the graph of its adjoint, which proves helpful in analyzing closures. On $\mathcal{H} \oplus \mathcal{H}$, consider the map

$$J(v, w) := (w, -v).$$

Clearly $J^2 = -I$, and it is easy to check that J is unitary.

Lemma 3.12. *For an operator T, the graph of T^* is the subspace*

$$\Gamma(T^*) = J(\Gamma(T)^\perp). \tag{3.12}$$

Proof By Definition 3.4, a pair (v, w) is contained in $\Gamma(T^*)$ if and only if

$$\langle v, Tu \rangle = \langle w, u \rangle \quad \text{for all } u \in \mathcal{D}(T).$$

Since (u, Tu) represents a general point in $\Gamma(T)$, this is equivalent to the condition that

$$(w, -v) \in \Gamma(T)^\perp.$$

In other words,

$$\Gamma(T^*) = J(\Gamma(T)^\perp).$$

□

Since orthogonal complements are closed, and closure is preserved under the unitary map J, an immediate consequence of (3.12) is the following:

Corollary 3.13. *The graph of the adjoint of an operator is closed.*

The formula (3.12) also allows us to clarify the issue noted in the previous section that the adjoint might not be densely defined.

Theorem 3.14. *An operator T is closable if and only if $\mathcal{D}(T^*)$ is dense, and in this case*

$$\overline{T} = T^{**}.$$

Proof Suppose $\mathcal{D}(T^*)$ is dense. Then T^{**} is well defined, and applying (3.12) to its graph yields

$$\Gamma(T^{**}) = J(\Gamma(T^*)^{\perp}). \tag{3.13}$$

A second application of (3.12), this time to $\Gamma(T^*)$ on the right-hand side of (3.13), gives

$$\begin{aligned}\Gamma(T^{**}) &= J(J\Gamma(T)^{\perp})^{\perp}\\ &= (\Gamma(T)^{\perp})^{\perp}\\ &= \overline{\Gamma(T)}.\end{aligned}$$

This shows that T is closable with $\overline{T} = T^{**}$.

To prove the converse statement, suppose that $\mathcal{D}(T^*)$ is not dense. By (2.27) this implies that $\mathcal{D}(T^*)^{\perp} \neq \{0\}$, so there exists a nonzero vector $v \in \mathcal{D}(T^*)^{\perp}$. Note that $(v, 0)$ is orthogonal to $\Gamma(T^*)$, which implies that

$$(0, -v) \in J(\Gamma(T^*)^{\perp}).$$

The set $J\Gamma(T^*)^{\perp}$ is equal to $\overline{\Gamma(T)}$ by (3.12) and (2.27), so $(0, -v) \in \overline{\Gamma(T)}$. Since $v \neq 0$, this shows that $\overline{\Gamma(T)}$ is not the graph of an operator, and therefore T is not closable. □

3.3.2 Closed Graph Theorem

We noted above that a bounded operator is trivially closed. The following theorem gives a partial converse statement, requiring the extra assumption that the domain is the full Hilbert space.

Theorem 3.15 (Closed Graph Theorem). *Suppose T is an operator on $\mathcal{H}$ with $\mathcal{D}(T) = \mathcal{H}$. Then T is bounded if and only if it is closed.*

Proof The forward implication follows by continuity, so we assume that T is closed and $\mathcal{D}(T) = \mathcal{H}$. Our first goal is to show that $\mathcal{D}(T^*)$ is closed.

Suppose that $\{v_n\} \subset \mathcal{D}(T^*)$ is a sequence such that $v_n \to v \in \mathcal{H}$. By the definition of the adjoint,

$$\langle T^* v_n, w\rangle = \langle v_n, Tw\rangle, \qquad \text{for all } w \in \mathcal{H}. \tag{3.14}$$

Consider $\{\langle T^* v_n, \cdot\rangle\}_{n=1}^{\infty}$ as a family of functionals in $\mathcal{H}'$. The family is pointwise bounded, because, for each $w \in \mathcal{H}$,

$$\begin{aligned}\sup_n |\langle T^* v_n, w\rangle| &= \sup_n |\langle v_n, Tw\rangle| \\ &\le \|Tw\| \sup_n \|v_n\| \\ &< \infty.\end{aligned}$$

Hence, by the uniform boundedness principle (Theorem 2.13),

$$M := \sup_n \|T^* v_n\| < \infty.$$

Since $v_n \to v$, it follows from (3.14) that

$$\begin{aligned}|\langle v, Tw\rangle| &\le \|w\| \limsup_{n\to\infty} \|T^* v_n\| \\ &\le M\|w\|\end{aligned}$$

for all $w \in \mathcal{H}$. Therefore $v \in \mathcal{D}(T^*)$ by (3.7), and hence $\mathcal{D}(T^*)$ is closed.

Since $\mathcal{D}(T^*)$ is also dense by Theorem 3.14, we have $\mathcal{D}(T^*) = \mathcal{H}$ and the relation

$$\langle Tv, w\rangle = \langle v, T^* w\rangle$$

holds for all $v, w \in \mathcal{H}$. Now consider the family of linear functionals $\{\langle Tv, \cdot\rangle : \|v\| = 1\}$. This family is pointwise bounded, because

$$\sup_{\|v\|=1} |\langle Tv, w\rangle| = \sup_{\|v\|=1} |\langle v, T^* w\rangle| \le \|T^* w\|,$$

for $w \in \mathcal{H}$. Therefore, by the uniform boundedness principle,

$$\|T\| := \sup_{\|v\|=1} \|Tv\| < \infty.$$

□

One straightforward corollary of the closed graph theorem is the Heiliger–Toeplitz theorem, which says that a self-adjoint operator with domain $\mathcal{H}$ is bounded. This result can also be derived directly from the uniform bounded principle, as seen in Exercise 2.9.

3.3.3 Invertibility

For bounded operators, the appropriate definition of invertibility is clear, because $\mathcal{L}(\mathcal{H})$ forms an algebra. We say that $T \in \mathcal{L}(\mathcal{H})$ is *invertible* if there exists $T^{-1} \in \mathcal{L}(\mathcal{H})$ such that $T^{-1}T = I$ and $TT^{-1} = I$. Because $\mathcal{L}(\mathcal{H})$ is associative, these conditions determine T^{-1} uniquely.

For unbounded operators, it makes logical sense to allow unbounded inverses. However, our applications will in fact require the inverse to be bounded, and there is no standard convention either way. To avoid potential confusion, we adopt the following terminology:

Definition 3.16. An operator T on $\mathcal{H}$ has a *bounded inverse* if there exists $T^{-1} \in \mathcal{L}(\mathcal{H})$ such that $TT^{-1} = I$ on $\mathcal{H}$ and $T^{-1}T = I$ on $\mathcal{D}(T)$.

One important application of the closed graph theorem is the fact (alluded to in Section 2.3) that for a bounded operator the existence of an inverse is sufficient to guarantee that the inverse is bounded.

Theorem 3.17 (Inverse Mapping Theorem). *An operator has a bounded inverse if and only if it is closed and bijective. In particular, for bounded operators invertibility is equivalent to bijectivity.*

Proof Assume that T has a bounded inverse T^{-1}. The graphs of T and T^{-1} are related by transposition, so T is closed because T^{-1} is closed. The conditions that $TT^{-1} = I$ on $\mathcal{H}$ and $T^{-1}T = I$ on $\mathcal{D}(T)$ imply that T is bijective as a map $\mathcal{D}(T) \to \mathcal{H}$.

Conversely, suppose that T is closed and bijective. Bijectivity implies that T^{-1} exists as a linear map $\mathcal{H} \to \mathcal{D}(T)$. As above, T^{-1} is closed because T is closed. Therefore T^{-1} is bounded by the closed graph theorem. □

The existence of a bounded inverse can be expressed in terms of norm estimates. We will say that the operator T is *bounded away from zero* if there exists a constant $c > 0$ such that

$$\|Tv\| \geq c\|v\| \tag{3.15}$$

for all $v \in \mathcal{D}(T)$.

Theorem 3.18. *An operator has a bounded inverse if and only if it is closed, bounded away from zero, and its range is dense in $\mathcal{H}$.*

This result is obvious in one direction: If T has a bounded inverse T^{-1}, then $\text{range}(T) = \mathcal{H}$ and the boundedness of T^{-1} implies

$$\|v\| \leq \|T^{-1}\| \|Tv\|$$

for all $v \in \mathcal{D}(T)$. For the proof of the converse statement see Exercise 3.4.

3.4 Symmetry and Self-adjointness

For matrices, the terms "symmetric" and "self-adjoint" are synonymous. A real symmetric matrix A satisfies $A^{\mathrm{t}} = A$, where the superscript denotes the transpose. In the complex case the corresponding symmetry condition is $\overline{A}^{\mathrm{t}} = A$, which is also called Hermitian. In either the real or complex case, the condition is that A is equal to its adjoint with respect to the Euclidean inner product.

In the theory of unbounded operators, we make a distinction between symmetry and self-adjointness. The latter term is given a literal interpretation:

Definition 3.19. A operator A is *self-adjoint* if $A^* = A$.

Note that since the domains of operators are implicit in the notation, self-adjointness requires that $\mathcal{D}(A^*) = \mathcal{D}(A)$, which means that $\mathcal{D}(A)$ must be chosen precisely so that the definition (3.7) of the adjoint domain reproduces the same space. A self-adjoint operator is closed, since the adjoint is closed by Corollary 3.13.

It is useful to take a more relaxed interpretation of symmetry, for which the choice of domain is not so rigid. An operator A is *symmetric* if

$$\langle Au, v\rangle = \langle u, Av\rangle \tag{3.16}$$

for all $u, v \in \mathcal{D}(A)$. Note that the symmetry of A implies that the functional $u \mapsto \langle v, Au\rangle$ is bounded for $v \in \mathcal{D}(A)$ because

$$\begin{aligned}|\langle v, Au\rangle| &= |\langle Av, u\rangle| \\ &\leq \|Av\|\|u\|,\end{aligned}$$

by (3.16) and Cauchy–Schwarz. Therefore, $\mathcal{D}(A) \subset \mathcal{D}(A^*)$. This implies that the symmetry property is equivalent to

$$A \subset A^*,$$

and shows in particular that a symmetric operator is closable.

An important class of symmetric operators is the set of positive operators. We say that A is *positive*, and write $A \geq 0$, if

$$\langle v, Av\rangle \geq 0$$

for all $v \in \mathcal{D}(A)$, i.e., the associated quadratic form is positive definite. A positive operator is necessarily symmetric; see Exercise 3.7 for the proof.

For differential operators, the properties of symmetry and positivity are closely related to integration by parts.

Example 3.20. Let $\Omega \subset \mathbb{R}^n$ be a bounded open set whose boundary $\partial\Omega$ is piecewise C^1. Green's identity gives the formula

$$\langle f, -\Delta f\rangle = \int_{\Omega} |\nabla f|^2 \, d^n x - \int_{\partial\Omega} \overline{f} \frac{\partial f}{\partial \nu} \, dS \tag{3.17}$$

for $f \in C^2(\overline{\Omega})$, where $\partial/\partial\nu$ denotes the outward normal derivative to $\partial\Omega$. Consider the classical (homogeneous) boundary conditions,

$$\text{Dirichlet: } f|_{\partial\Omega} = 0, \qquad \text{Neumann: } \frac{\partial f}{\partial \nu}\Big|_{\partial\Omega} = 0.$$

Either of these will imply that the integral over $\partial\Omega$ in (3.17) vanishes, leaving $\langle f, -\Delta f\rangle \geq 0$. The (negative) Laplacian is thus a positive operator (and hence symmetric) on a domain consisting $C^2(\overline{\Omega})$ restricted by either Dirichlet or Neumann boundary conditions.

It is easy to see that the Laplacian is not closed on a domain defined as a subspace of $C^2(\overline{\Omega})$. Hence the resulting operators are not self-adjoint. (We will develop self-adjoint extensions corresponding to the classical boundary conditions in Chapter 6.) ♢

Symmetry alone is not sufficient as a hypothesis for any of the major results of spectral theory. It should thus be thought of as an intermediate property which allows us to separate the relatively easy problem of establishing "formal" self-adjointness, meaning (3.16), from the more difficult task of finding the exact domain.

3.4.1 *Self-adjoint Operators*

Let us first consider some operators we have previously encountered, where self-adjointness is clear.

Example 3.21. On $L^2[0,1]$ consider the operator $T := -i\frac{d}{dx}$, with domain given by

$$\mathcal{D}(T) = \left\{ f \in AC[0,1] : \; f(0) = f(1), \; f' \in L^2[0,1] \right\}.$$

For $f, g \in AC[0,1]$, integration by parts gives

$$-i \int_0^1 \overline{f} g' \, dt = -i \overline{f} g \Big|_0^1 + i \int_0^1 \overline{f'} g \, dt.$$

The boundary term vanishes for $f, g \in \mathcal{D}(T)$ because of the periodic boundary condition, implying that

$$\langle f, Tg\rangle = \langle Tf, g\rangle,$$

so T is symmetric.

We claim that T is self-adjoint. To see this, we must show that $\mathcal{D}(T^*) \subset \mathcal{D}(T)$. Suppose $f \in \mathcal{D}(T^*)$, and set $g = T^* f$. As in Example 3.6, we set

$$h(x) := \int_0^x g(t)\, dt,$$

and use integration by parts to derive

$$\langle g, \phi \rangle + \langle h, \phi' \rangle = \overline{h(1)}\phi(1),$$

for $\phi \in \mathcal{D}(A)$. Since $g = T^* f$, this translates to

$$\langle f - ih, T\phi \rangle = \overline{h(1)}\phi(1).$$

Setting $\phi \equiv 1$ implies that $h(1) = 0$, and thus

$$f - ih \in \text{range}(T)^{\perp}. \tag{3.18}$$

Suppose that $u \in C^1[0, 1]$ with

$$\int_0^1 u\, dt = 0.$$

Then $u = Tv$ for $v \in \mathcal{D}(A)$ defined by

$$v(x) := -i \int_0^x u(t)\, dt.$$

Hence $C^1[0, 1] \cap \{1\}^{\perp} \subset \text{range}(T)$, which implies that $\text{range}(T)^{\perp}$ is the one-dimensional space of constant functions. It thus follows from (3.18) that

$$f(x) = f(0) + ih(x).$$

This implies that $f \in AC[0, 1]$, and hence $f \in \mathcal{D}(T)$. ♢

Example 3.22. In Example 3.5 we noted that the adjoint of a multiplication operator M_f on $L^2(X, d\mu)$ is given by

$$M_f^* = M_{\overline{f}},$$

with the same domain $\mathcal{D}(M_f)$. Therefore M_f is self-adjoint if and only if f is real-valued almost everywhere. ♢

Example 3.23. As we noted in Example 3.3, the Laplacian $-\Delta$ is unitarily equivalent to $M_{|\xi|^2}$ via the Fourier transform. Example 3.22 shows that $M_{|\xi|^2}$ is self-adjoint, and we saw in Example 3.3 that the domain $\mathcal{D}(M_{|\xi|^2})$ is mapped back

to $H^2(\mathbb{R}^n)$ under the Fourier transform. Therefore, $-\Delta$ is self-adjoint on $L^2(\mathbb{R}^n)$ with $\mathcal{D}(-\Delta) = H^2(\mathbb{R}^n)$. $\Diamond$

These cases are very special, in that the exact domain for self-adjointness is relatively easy to determine. Example 3.20 illustrates the more common situation, where a natural choice of boundary conditions yields a symmetric operator that is not self-adjoint. We are then faced with the problem of finding a self-adjoint extension of the symmetric operator. In general, neither existence nor uniqueness of such an extension is guaranteed.

Since a symmetric operator is closable, one obvious candidate for a self-adjoint extension is the closure itself. The class of symmetric operators for which this works is distinguished in the following:

Definition 3.24. A symmetric operator is *essentially self-adjoint* if its closure is self-adjoint.

It is not difficult to see that A is essentially self-adjoint if and only if

$$\overline{A} = A^*. \tag{3.19}$$

We leave the proof to Exercise 3.9. A domain on which a symmetric operator is essentially self-adjoint is called a *core domain*. The core is not unique, but the following result shows that the extension is independent of the choice of core.

Lemma 3.25. *For an essentially self-adjoint operator, the closure is the unique self-adjoint extension.*

Proof Suppose that A is essentially self-adjoint and B is a self-adjoint extension of A. The closure $\overline{A}$ is the smallest closed extension of A, so $\overline{A} \subset B$. Taking the adjoint of this relation gives $B \subset \overline{A}$, because both $\overline{A}$ and B are self-adjoint. Therefore $B = \overline{A}$. $\square$

Example 3.26. Consider the Laplacian $-\Delta$ acting on $L^2(\mathbb{R}^n)$, with domain equal to $\mathcal{D}(A) = C_0^\infty(\mathbb{R}^n)$. The domain $\mathcal{D}(-\Delta^*)$ then consists of functions $f \in L^2(\mathbb{R}^n)$ such that the map

$$\phi \mapsto \langle f, -\Delta\phi \rangle$$

is bounded for $\phi \in C_0^\infty(\mathbb{R}^n)$. Since

$$\langle f, -\Delta\phi \rangle = \langle \hat{f}, |\xi|^2 \hat{\phi} \rangle,$$

by the Plancherel theorem, $f \in \mathcal{D}(-\Delta^*)$ if and only if there exists a function $g \in L^2(\mathbb{R}^n)$ such that $g - |\xi|^2 \hat{f}$ is orthogonal to $\mathcal{F}(C_0^\infty(\mathbb{R}^n))$. This implies that $g = |\xi|^2 \hat{f}$, because the space $C_0^\infty(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$ and the Fourier transform is unitary.

This argument shows that $f \in \mathcal{D}(-\Delta^*)$ if and only if $|\xi|^2 \hat{f} \in L^2(\mathbb{R}^n)$. In other words,

$$\mathcal{D}(A^*) = H^2(\mathbb{R}^n),$$

where $H^2(\mathbb{R}^n)$ is the Sobolev space introduced in Example 3.23. Since $-\Delta$ is self-adjoint on the domain $H^2(\mathbb{R}^n)$, we conclude that $-\Delta$ is essentially self-adjoint on the core $C_0^\infty(\mathbb{R}^n)$. ♢

In Chapter 7 we will study Schrödinger operators, which have the form $-\Delta + V$, with V a potential function on $\mathbb{R}^n$ acting as a multiplication operator. We can extend the observation from Example 3.26 to certain Schrödinger operators using the following:

Lemma 3.27. *If A and B are self-adjoint operators and B is bounded, then $A + B$ is self-adjoint with domain equal to $\mathcal{D}(A)$. Furthermore, $A + B$ is essentially self-adjoint on a core domain for A.*

Proof The sum $T := A + B$ is clearly symmetric on the domain $\mathcal{D}(T) := \mathcal{D}(A)$. Suppose that $u \in \mathcal{D}(T^*)$, which means that the functional

$$v \mapsto \langle u, Tv\rangle$$

is bounded on $\mathcal{D}(T)$. By writing

$$\langle u, Av\rangle = \langle u, Tv\rangle - \langle u, Bv\rangle,$$

and using the fact that B is bounded, we can deduce that $u \in \mathcal{D}(A^*)$ also. If A is self-adjoint, then this shows that $\mathcal{D}(T^*) \subset \mathcal{D}(A) = \mathcal{D}(T)$. Hence T is self-adjoint.

Suppose now that A is essentially self-adjoint on $\mathcal{D}(A)$. If $u \in \mathcal{D}(\overline{A})$, then there exists a sequence $\{u_n\}$ such that $u_n \to u$ and Au_n converges to $\overline{A}u$. Since B is bounded, Bu_n also converges to Bu. Therefore the sequence Tu_n converges to $Au + Bu$. This shows that $u \in \mathcal{D}(\overline{T})$ with $\overline{T}u = \overline{A}u + Bu$. Similarly, we can argue that $u \in \mathcal{D}(\overline{T})$ implies that $u \in \mathcal{D}(\overline{A})$ with $\overline{A}u = \overline{T}u - Bu$. We can conclude that $\mathcal{D}(\overline{T}) = \mathcal{D}(\overline{A})$ and

$$\overline{T} = \overline{A} + B.$$

Since $\overline{A}$ is self-adjoint by assumption, the self-adjointness of $\overline{T}$ follows from the first part of the proof. □

Example 3.28. Suppose $V \in L^\infty(\mathbb{R}^n)$ is real-valued. Then the multiplication operator M_V is bounded and self-adjoint on $L^2(\mathbb{R}^n)$ by Examples 3.2 and 3.22. Therefore, by Lemma 3.27 and Example 3.26, the Schrödinger operator $-\Delta + V$ is self-adjoint on $H^2(\mathbb{R}^n)$ and essentially self-adjoint on $C_0^\infty(\mathbb{R}^n)$. ♢

In Section 7.2, we will improve the result of Lemma 3.27 by weakening the restriction to allow B to be unbounded.

3.4.2 Criteria for Self-adjointness

So far, the only means that we have to verify that a symmetric operator is self-adjoint is to explicitly check its domain against the definition (3.7) of the adjoint domain. Although we have given a few examples in Section 3.4.1 where this works, direct calculation of the adjoint domain is usually difficult. In this section we will develop some alternative criteria for self-adjointness and essential self-adjointness. Many applications of these criteria will be seen in later sections.

Consider a symmetric operator A. For $v \in \mathcal{D}(A)$, the symmetry condition (3.16) implies that

$$\operatorname{Im}\langle v, (A - z)v\rangle = -(\operatorname{Im} z)\|v\|^2$$

for $z \in \mathbb{C}$. By an application of the Cauchy–Schwarz inequality, this implies

$$\|v\| \leq \frac{\|(A - z)v\|}{|\operatorname{Im} z|} \tag{3.20}$$

for $v \in \mathcal{D}(A)$. This is the crucial observation behind the following:

Theorem 3.29. *Suppose that A is a symmetric operator, and let $z \in \mathbb{C}$ be strictly complex. The following conditions are equivalent:*

(a) *A is self-adjoint.*
(b) *A is closed and both $A^* - z$ and $A^* - \bar{z}$ are injective.*
(c) *Both $A - z$ and $A - \bar{z}$ have bounded inverses.*
(d) *Both $A - z$ and $A - \bar{z}$ are surjective.*

Proof Assume that A is self-adjoint and therefore closed. The operators $A \pm z$ are injective by (3.20), because A is assumed symmetric. Thus (a) implies (b).

Now assume (b). By Lemma 3.7,

$$\ker(A^* - \bar{z}) = \operatorname{range}(A - z)^{\perp}. \tag{3.21}$$

Hence the assumption that $A^* - \bar{z}$ is injective implies that $\operatorname{range}(A - z)$ is dense. Since $A - z$ is closed by assumption, and bounded away from zero by (3.20), Theorem 3.18 shows that $A - z$ has a bounded inverse. Similarly, the assumption that $A^* - z$ is injective implies $A - \bar{z}$ has a bounded inverse. This shows that (b) implies (c).

It is trivial that (c) implies (d), so it remains to prove that (d) implies (a). Assume that $A - z$ and $A - \bar{z}$ are surjective. To prove self-adjointness it suffices to show $\mathcal{D}(A^*) \subset \mathcal{D}(A)$. Let $u \in \mathcal{D}(A^*)$. The surjectivity of $A - z$ implies that there exists $v \in \mathcal{D}(A)$ such that

$$(A - z)v = (A^* - z)u.$$

Since $Av = A^*v$ by symmetry, subtracting one side from the other gives

$$(A^* - z)(u - v) = 0. \tag{3.22}$$

The surjectivity of $A - \bar{z}$ implies that $A^* - z$ is injective by (3.21), so (3.22) shows that $u = v$ and hence $u \in \mathcal{D}(A)$. Thus $\mathcal{D}(A^*) \subset \mathcal{D}(A)$, proving that A is self-adjoint. □

With a fairly straightforward adaptation of the proof of Theorem 3.29, we can produce a corresponding result for essential self-adjointness.

Theorem 3.30. *Suppose that A is a symmetric operator, and let $z \in \mathbb{C}$ be strictly complex. The following conditions are equivalent:*

(a) *A is essentially self-adjoint.*
(b) *Both $A^* - z$ and $A^* - \bar{z}$ are injective.*
(c) *Both $A - z$ and $A - \bar{z}$ have dense range.*

Proof If A is essentially self-adjoint, then A^* is self-adjoint by (3.19). Thus (b) follows from (a) by Theorem 3.29. And (c) follows immediately from (b) by (3.21).

It remains to show that (c) implies (a). Assume that range$(A - z)$ is dense. Clearly this also implies that range$(\overline{A} - z)$ is dense. Furthermore, a simple argument (Exercise 3.8) shows that $\overline{A}$ is symmetric. Thus $\overline{A} - z$ is bounded away from zero by (3.20). It now follows that $\overline{A} - z$ has a bounded inverse by Theorem 3.18. The same reasoning applies to $\overline{A} - \bar{z}$, and Theorem 3.29 then shows that $\overline{A}$ is self-adjoint. □

Theorems 3.29 and 3.30 are perhaps the best general tools we have for defining self-adjoint operators. We will see ample demonstration of this in later chapters. For now we will illustrate their usage in some simple one-dimensional cases.

Example 3.31. On $L^2(0, 1)$, consider the operator

$$L_0 = -\Delta, \qquad \mathcal{D}(L_0) = C_0^\infty(0, 1).$$

Clearly L_0 is symmetric, since integration by parts on $C_0^\infty(0, 1)$ produces no boundary terms, but it is not essentially self-adjoint. To show this, by Theorem 3.30 it suffices to produce a nonzero element of $L^2(0, 1)$ which is orthogonal to the range of $L_0 - z$ for z strictly complex.

For $u \in C^\infty[0, 1]$ and $\phi \in C_0^\infty(0, 1)$, we have

$$\langle u, (L_0 - z)\phi\rangle = \langle -u'' - \bar{z}u, \phi\rangle. \tag{3.23}$$

The equation $u'' + \bar{z}u = 0$ admits nontrivial solutions, such as $u(x) = e^{i\sqrt{\bar{z}}x}$. By (3.23) these solutions are orthogonal to the range of $L_0 - z$. Therefore L_0 is not essentially self-adjoint. ♢

Example 3.32. The *Robin boundary conditions* on $(0, 1)$ are given by

$$a_0u(0) - b_0u'(0) = 0, \qquad a_1u(1) + b_1u'(1) = 0, \tag{3.24}$$

where the coefficients a_j, b_j are real and $(a_j, b_j) \neq (0, 0)$ for $j = 0, 1$. These are a generalization of the classical conditions introduced in Example 3.20, reducing to Dirichlet if $b_j = 0$ and Neumann if $a_j = 0$.

Let $L = -\Delta$ on $L^2(0, 1)$, with domain $\mathcal{D}(L)$ consisting of smooth functions on $[0, 1]$ satisfying (3.24). For $f, g \in \mathcal{D}(L)$,

$$\langle f, -\Delta g\rangle - \langle -\Delta f, g\rangle = (f'g - g'f)\Big|_0^1$$

The vectors (f, f') and (g, g') are linearly dependent at 0 and 1, by (3.24), so $f'g - g'f$ vanishes at the endpoints. Therefore L is symmetric.

To see that L is essentially self-adjoint, we can use the ODE technique of variation of parameters to construct the integral kernel for $(L - z)^{-1}$. For $j = 0, 1$, let w_j be a nonzero solution of

$$w_j'' + zw_j = 0, \qquad \left[a_j w_j + (-1)^{j+1} b_j w_j'\right]\Big|_{x=j} = 0,$$

where z is strictly complex. Each w_j is a linear combination of the functions $e^{\pm i\sqrt{z}x}$, and is therefore smooth. We claim that the pair w_0, w_1 is linearly independent. Otherwise, w_0 would satisfy (3.24) at both endpoints. Then, by symmetry,

$$0 = \langle w_0, w_0''\rangle - \langle w_0'', w_0\rangle = (\bar{z} - z)\|w_0\|^2.$$

Since z is strictly complex, this is not possible for w_0 nonzero.

For solutions of a second order ODE, the Wronskian,

$$W[w_0, w_1] := w_0 w_1' - w_0' w_1,$$

is given by a constant c_0. Linear independence implies that $c_0 \neq 0$, allowing us to define the integral kernel,

$$K(x, y) := \frac{1}{c_0}\begin{cases} w_0(x)w_1(y), & x \leq y, \\ w_0(y)w_1(x), & x \geq y. \end{cases}$$

For $f \in C^\infty[0, 1]$, a simple calculation with the fundamental theorem of calculus then shows that

$$u(x) := \int_0^1 K(x, y)f(y)\,dy$$

yields a solution of $u'' + zu = f$ with $u \in \mathcal{D}(L)$. Hence f lies in the range of $L - z$. Since $C^\infty[0, 1]$ is dense in $L^2(0, 1)$, this shows that $L - z$ has dense range. The same construction applies to $L - \bar{z}$, so we conclude that L is essentially self-adjoint by Theorem 3.30. ♢

In Example 3.32 we can see the dramatic effect that the choice of extension could have on the properties of the operator. For any choice of parameters a_0, b_0 and a_1, b_1, the Robin Laplacian L is an extension of the positive operator $-\Delta$ with domain $C_0^\infty(0, 1)$. Despite the positivity of $-\Delta$ on this smaller domain, not all of the extensions are positive. Indeed, if we take $a_0 = a_1 = \alpha$, $b_0 = -1$, and $b_1 = 1$, then $\phi(x) := e^{-\alpha x}$ lies in $\mathcal{D}(L)$ and satisfies $L\phi = -\alpha^2\phi$. This shows that we can choose boundary conditions for which the infimum of $\langle \phi, L\phi\rangle / \|\phi\|^2$ is arbitrarily negative.

3.4.3 Friedrichs Extension

For a positive symmetric operator, we usually look for a self-adjoint extension which is also positive. The remark at the end of the previous section raises the question of whether or not this is possible. The answer is clear in the essentially self-adjoint case, because the closure of a positive operator is easily seen to be positive. In the general case it is not so obvious, but positive extensions do in fact exist in general.

Kurt Friedrichs solved this problem by developing an extension method based on the quadratic form $u \mapsto \langle u, Au\rangle$ associated with a positive operator A. Note that we could always shift a positive operator by a constant, without changing the domain of the extension. Therefore, the natural context for the Friedrichs method is a symmetric operator S which is *semi-bounded*, meaning that there exists a constant $a \in \mathbb{R}$ such that

$$\langle u, Su\rangle \geq a\|u\|^2 \tag{3.25}$$

for all $u \in \mathcal{H}$.

Theorem 3.33 (Friedrichs Extension). *Suppose that S is a symmetric operator on $\mathcal{H}$ which is semi-bounded in the sense of* (3.25)*. Then there exists a self-adjoint extension of S satisfying the same bound.*

Proof By adding a constant to S, if necessary, it suffices to consider the case $a = 1$, where S satisfies

$$\langle u, Su\rangle \geq \|u\|^2. \tag{3.26}$$

Since S is symmetric, (3.26) implies that the sesquilinear form,

$$Q[u, v] := \langle u, Sv\rangle, \tag{3.27}$$

defines an inner product on $\mathcal{D}(S)$. Let $\mathcal{H}_Q$ denote the abstract Hilbert space completion of $\mathcal{D}(S)$ with respect to $Q[\cdot, \cdot]$, as described in Section 2.4. Elements of $\mathcal{H}_Q$ are equivalence classes of sequences in $\mathcal{D}(S)$ which are Cauchy with respect to the norm $\|\cdot\|_Q$ associated with $Q[\cdot, \cdot]$.

Our goal is to identify $\mathcal{H}_Q$ with a subspace of $\mathcal{H}$. We first note that the inclusion of $(\mathcal{D}(S), Q[\cdot,\cdot])$ in $\mathcal{H}$ given by the identity map is continuous, by (3.26). It therefore admits a unique extension to a continuous linear map $J : \mathcal{H}_Q \to \mathcal{H}$.

We claim that J is injective. To prove this, suppose that $w \in \mathcal{H}_Q$ and $J(w) = 0$. Since $\mathcal{D}(S)$ is embedded in $\mathcal{H}_Q$, we can find a sequence $\{v_k\} \subset \mathcal{D}(S)$ such that $v_k \to w$ in $\mathcal{H}_Q$. By continuity, $J(v_k) \to J(w)$ in $\mathcal{H}$ as $k \to \infty$. Since $J(w) = 0$, and J reduces to the identity on $\mathcal{D}(S)$, this implies that

$$\lim_{k\to\infty} \|v_k\| = 0. \tag{3.28}$$

Using the fact that the inner product on $\mathcal{H}_Q$ is a continuous extension of (3.27), for $u \in \mathcal{D}(S)$ we can compute

$$\begin{aligned}\langle w, u\rangle_{\mathcal{H}_Q} &= \lim_{k\to\infty} Q[v_k, u] \\ &= \lim_{k\to\infty} \langle v_k, Su\rangle.\end{aligned}$$

Hence, (3.28) implies that $\langle w, u\rangle_{\mathcal{H}_Q} = 0$ for all $u \in \mathcal{D}(S)$. Since $\mathcal{D}(S)$ is dense in $\mathcal{H}_Q$, it follows that $w = 0$.

Using the fact that J is injective, we can interpret $\mathcal{H}_Q$ as a subspace of $\mathcal{H}$ by identifying $w \in \mathcal{H}_Q$ with $J(w) \in \mathcal{H}$. The desired self-adjoint extension A can now be defined on the domain

$$\mathcal{D}(A) := \left\{u \in \mathcal{H}_Q : \left|\langle u, v\rangle_{\mathcal{H}_Q}\right| \le C\|v\| \text{ for all } v \in \mathcal{H}_Q\right\}. \tag{3.29}$$

The condition on u in (3.29) implies that $v \mapsto \langle u, v\rangle_{\mathcal{H}_Q}$ extends to a bounded linear functional on $\mathcal{H}$. The Riesz lemma then defines a unique element Au such that

$$\langle u, v\rangle_{\mathcal{H}_Q} = \langle Au, v\rangle, \tag{3.30}$$

for all $v \in \mathcal{H}_Q$. It is easy to check that $u \mapsto Au$ is linear, and (3.30) implies that A is symmetric. Moreover, (3.27) implies that $S \subset A$.

To see that A is self-adjoint, we use a variant of the argument from Theorem 3.29. For $u \in \mathcal{H}$, the functional $\langle u, \cdot\rangle$ is continuous on $\mathcal{H}_Q$, by (3.26). Since $\mathcal{H}_Q$ is a Hilbert space, the Riesz lemma gives $w \in \mathcal{H}_Q$ such that

$$\langle u, v\rangle = \langle w, v\rangle_{\mathcal{H}_Q}$$

for all $v \in \mathcal{H}_Q$. By (3.29), we have $w \in \mathcal{D}(A)$ and $Aw = u$. Since u was arbitrary, this proves that A is surjective. By Lemma 3.7, A^* is therefore injective.

Now consider $u \in \mathcal{D}(A^*)$. By the surjectivity of A, we have $A^*u = Aw$ for some $w \in \mathcal{D}(A)$. Since $A \subset A^*$, this implies $A^*(u - w) = 0$. Hence, $u = w$, because A^* is injective, which proves that $u \in \mathcal{D}(A)$. This establishes the inclusion $\mathcal{D}(A^*) \subset \mathcal{D}(A)$, proving that A is self-adjoint.

By (3.30), and the fact that $\langle\cdot,\cdot\rangle_{\mathcal{H}_Q}$ is a continuous extension of $Q[\cdot,\cdot]$,

$$\langle u, Au\rangle \geq \|u\|^2$$

for all $u \in \mathcal{D}(A)$. □

The Friedrichs construction can be applied more generally to construct a self-adjoint operator from a quadratic form defined on a dense subspace $\mathcal{H}_1 \subset \mathcal{H}$. The requirements on $\mathcal{H}_1$ for this construction are evident in the proof of Theorem 3.33. We need for $\mathcal{H}_1$ to be a Hilbert space with respect to $\langle\cdot,\cdot\rangle_{\mathcal{H}_1}$ and for the embedding $\mathcal{H}_1 \to \mathcal{H}$ to be continuous, meaning that there exists $C > 0$ such that

$$\|u\| \leq C\|u\|_{\mathcal{H}_1}$$

for all $u \in \mathcal{H}_1$. We will see some important applications of the Friedrichs extension in Section 6.1.

3.5 Compact Operators

In 1916 Riesz observed that many qualitative features of the spectrum of integral operators could be deduced from a property that he referred to as "complete continuity." The modern formulation of this condition is the following:

Definition 3.34. An operator $T \in \mathcal{L}(\mathcal{H})$ is *compact* if for any bounded sequence $\{u_k\} \subset \mathcal{H}$, the sequence $\{Tu_k\}$ has a convergent subsequence.

Another way to say this is that T maps bounded subsets to relatively compact subsets. The same definition applies also to maps between Banach spaces, but our focus is on Hilbert space operators.

The most basic example of a compact operator is a bounded operator of finite rank. If $\{Tu_k\}$ is a bounded sequence contained in a finite-dimensional subspace of $\mathcal{H}$, then the Bolzano–Weierstrass theorem implies that $\{Tu_k\}$ has a convergent subsequence. Thus all bounded finite-rank operators are compact.

It follows directly from Definition 3.34 that the sum of compact operators is compact. Moreover, if B is a bounded operator and T is compact, then the continuity of B implies that both BT and TB are compact. We can summarize these observations as follows:

Lemma 3.35. *The set of compact operators on $\mathcal{H}$ is an algebra and a two-sided ideal within $\mathcal{L}(\mathcal{H})$.*

Alongside these algebraic properties, we have an important analytic result, that the compact operators form a closed subspace with respect to the operator topology.

Theorem 3.36. *If $\{T_n\}$ is a sequence of compact operators on $\mathcal{H}$ that converges (with respect to the operator norm) to a bounded operator T, then T is compact.*

Proof Suppose $\{u_k\} \subset \mathcal{H}$ is a bounded sequence and let $M := \sup \|u_k\|$. Using the fact that T_1 is compact, we can choose a subsequence $\{u_{1,k}\}$ for which $\{T_1 u_{1,k}\}$ is convergent. Then we can take $\{u_{2,k}\} \subset \{u_{1,k}\}$ so that $\{T_2 u_{2,k}\}$ converges, and so on. The result is a nested set of subsequences $\{u_{m,k}\}$ for which $\{T_j u_{m,k}\}_{k=1}^{\infty}$ is convergent as $k \to \infty$ for $m \geq j$.

Now consider the diagonal subsequence $w_k := u_{k,k}$. We claim that the image sequence $\{T w_k\}$ is Cauchy. To show this, pick $\varepsilon > 0$. There exists an n such that

$$\|T_n - T\| \leq \varepsilon$$

by the norm convergence $T_n \to T$. For this value of n,

$$\|T w_k - T w_m\| \leq 2M\varepsilon + \|T_n w_k - T_n w_m\| \tag{3.31}$$

by the triangle inequality. The sequence $\{T_j w_k\}_{k=1}^{\infty}$ is convergent by the diagonal construction, and therefore Cauchy as well. Hence there exists $N > 0$ such that for $k, n \geq N$

$$\|T_n w_k - T_n w_m\| \leq \varepsilon.$$

Applying this to (3.31) gives

$$\|T w_k - T w_m\| \leq (2M + 1)\varepsilon,$$

for $k, n \geq N$. Since M is fixed and ε is arbitrary, we conclude that $\{T w_k\}$ is Cauchy, and hence convergent. □

The close connection between compact and finite-rank operators is demonstrated by the following result. Our proof is limited to separable Hilbert spaces for the sake of brevity, but the result does extend to the non-separable case.

Theorem 3.37. *Let $\mathcal{H}$ be a separable Hilbert space. An operator $T \in \mathcal{L}(\mathcal{H})$ is compact if and only if there exists a sequence of bounded finite-rank operators $\{T_n\}$ such that $\|T_n - T\| \to 0$.*

Proof Since bounded finite-rank operators are compact, limits of sequences of such operators with respect to the operator-norm topology are compact by Theorem 3.36.

To prove the converse, suppose that T is a compact operator. Our goal is to construct an approximating sequence of finite-rank operators. This is trivial if $\dim \mathcal{H} < \infty$ so we assume that $\dim \mathcal{H} = \infty$.

By Theorem 2.33, there exists an orthonormal basis $\{e_j\}_{j=1}^{\infty}$ for $\mathcal{H}$. Let P_n be the orthogonal projection onto the span of $\{e_1, \ldots, e_n\}$ and R_n the orthogonal projection onto $\{e_1, \ldots, e_n\}^{\perp}$, so that $P_n + R_n = I$. Clearly $P_n T$ is bounded and has finite-rank, so we can complete the proof by showing that $P_n T \to T$.

For $v \in \mathcal{H}$

$$\|R_n T v\|^2 = \sum_{j=n+1}^{\infty} \left|\langle e_j, Tv\rangle\right|^2.$$

Taking the supremum of both sides over $\|v\| = 1$ shows that $\|R_n T\|$ is decreasing as $n \to \infty$. This means that either $\|R_n T\| \to 0$ or $\|R_n T\| > \varepsilon$ for all n for some $\varepsilon > 0$. Assume that the latter holds. Then for each n we can choose $v_n \in \mathcal{H}$ some $\|v_n\| = 1$ such that

$$\|R_n T v_n\| \geq \varepsilon.$$

By the compactness of T, the sequence $\{T v_n\}$ admits a convergent subsequence: $T v_{n_k} \to w \in \mathcal{H}$ as $k \to \infty$. Then

$$\|R_{n_k} T v_{n_k}\| \leq \|R_{n_k} w\| + \|R_{n_k}(T v_{n_k} - w)\|,$$

and taking $k \to \infty$ gives

$$\varepsilon \leq \liminf_{k\to\infty} \|R_{n_k} w\|.$$

This is a contradiction, because

$$\|R_{n_k} w\|^2 = \sum_{j>n_k} \left|\langle e_j, w\rangle\right|^2,$$

and so $\|R_{n_k} w\| \to 0$ by Bessel's inequality. We conclude that $\|R_n T\| \to 0$, which proves that $P_n T \to T$. □

Using Theorem 3.37, we can develop an equivalent characterization of compactness, which is sometimes taken as the definition.

Theorem 3.38. *A bounded operator $T \in \mathcal{L}(\mathcal{H})$ is compact if and only if T maps weakly convergent sequences to convergent sequences.*

Proof Suppose first that T is compact and $\{u_k\} \subset \mathcal{H}$ is a sequence which converges to w in the weak sense. For each $v \in \mathcal{H}$, weak convergence implies a bound

$$\sup_k |\langle u_k, v\rangle| < \infty.$$

Thus, by uniform boundedness (Theorem 2.13), there exists $M > 0$ such that

$$\|u_k\| \leq M$$

for all k.

Given $\varepsilon > 0$, Theorem 3.37 implies that there exists a bounded finite-rank operator T_n such that $\|T_n - T\| < \varepsilon$. Replacing T by T_n thus gives

$$\|T(u_k - w)\| \leq \|T_n(u_k - w)\| + \varepsilon(M + \|w\|). \tag{3.32}$$

Since T_n is bounded and has finite rank, weak weak convergence $u_k \to w$ implies that $T_n(u_k - w) \to 0$ in norm. Thus taking $k \to \infty$ in (3.32) gives

$$\limsup_{k\to\infty} \|T(u_k - w)\| \leq \varepsilon(M + \|w\|).$$

This holds for all $\varepsilon > 0$, which proves that $Tu_k \to Tw$.

To prove the other direction, we assume that $T \in \mathcal{L}(\mathcal{H})$ maps weakly convergent sequences to strongly convergent sequences, and let $\{u_k\}$ be a bounded sequence in $\mathcal{H}$. Alaoglu's theorem (Theorem 2.37) gives a subsequence $\{u_{k_j}\}$ that converges weakly to some $w \in \mathcal{H}$. By assumption, the fact that $u_{k_j} - w \to 0$ in the weak sense implies the norm convergence $Tu_{k_j} \to Tw$. □

3.5.1 Hilbert–Schmidt Operators

The first compact operators to be studied were integral operators of a specific type, introduced by David Hilbert and Erhard Schmidt in 1907. Suppose that $(X, \mathcal{M}, \mu)$ is a measure space such that $L^2(X, d\mu)$ is separable (for example, a subset $\Omega \subset \mathbb{R}^n$ with Lebesgue measure). A *Hilbert–Schmidt operator* on $L^2(X, d\mu)$ is an integral operator defined by

$$Tf(x) := \int_X K(x, y) f(y)\, d\mu(y), \tag{3.33}$$

where the integral kernel satisfies

$$K \in L^2(X \times X, d\mu \otimes d\mu).$$

The integral on the right-hand side of (3.33) is defined for almost every x. To see this, note that

$$\int_X |K(x, y) f(y)|\, d\mu(y) \leq \|K(x, \cdot)\|_{L^2(X)} \|f\| \tag{3.34}$$

by the Cauchy–Schwarz inequality. The function $|K(\cdot, \cdot)|^2$ is integrable on $X \times X$, so $\|K(x, \cdot)\|_{L^2(X)} < \infty$ for almost every x by Fubini's theorem.

Theorem 3.39. *If T is a Hilbert–Schmidt operator with integral kernel K, then*

$$\|T\| \leq \|K\|_{L^2(X\times X)},$$

and T is compact.

Proof The operator-norm estimate follows from squaring (3.34) and integrating:

$$\begin{aligned}\|Tf\|^2 &= \int_X \left(\int_X |K(x,y)f(y)|\, d\mu(y) \right)^2 d\mu(x) \\ &\leq \int_X \|K(x,\cdot)\|^2_{L^2(X)} \|f\|^2 \, d\mu(x) \\ &= \|K\|^2_{L^2(X\times X)} \|f\|^2. \end{aligned} \tag{3.35}$$

By the assumption that $L^2(X, d\mu)$ is separable, there exists an orthonormal basis $\{\phi_k\}_{k=1}^\infty$. It is easy to check that $\{\phi_j \otimes \overline{\phi_k}\}_{j,k=1}^\infty$ gives a corresponding basis for $L^2(X \times X, d\mu \otimes d\mu)$. In terms of this basis, K has the expansion

$$K(x,y) = \sum_{j,k=1}^\infty a_{jk}\phi_j(x)\overline{\phi_k(y)}, \tag{3.36}$$

converging in the L^2 sense, where

$$a_{jk} := \langle \phi_j, T\phi_k \rangle. \tag{3.37}$$

By Theorem 2.34,

$$\|K\|^2_{L^2(X\times X)} = \sum_{j,k=1}^\infty |a_{jk}|^2. \tag{3.38}$$

The basis expansion (3.36) provides a natural bounded finite-rank approximation T_N, with integral kernel

$$K_N(x,y) := \sum_{j,k=1}^N a_{jk}\phi_j(x)\overline{\phi_k(y)}.$$

By the estimate used in (3.35),

$$\begin{aligned}\|T - T_N\|^2 &\leq \|K - K_N\|^2_{L^2(X\times X)} \\ &= \sum_{j,k>N} |a_{jk}|^2. \end{aligned}$$

By (3.38), this is the tail of a convergent series, implying that $\|T - T_N\| \to 0$ as $N \to \infty$. □

Example 3.40. On $L^2(0,1)$, consider the integration operator T, given by

$$Tf(x) := \int_0^x f(y)\,dy.$$

This can be written as an integral operator with the kernel,

$$K(x,y) = \begin{cases} 1, & x \le y, \\ 0, & x > y. \end{cases}$$

Clearly $K \in L^2((0,1)\times(0,1))$, so T is a Hilbert–Schmidt operator and therefore compact. ◊

If T has an integral kernel K, then the L^2 norm of K is called the Hilbert–Schmidt norm of T. The formula (3.38) allows us to extend this notion to operators on an abstract (separable) Hilbert space $\mathcal{H}$. We say that $T \in \mathcal{L}(\mathcal{H})$ is Hilbert–Schmidt if, for an orthonormal basis $\{\phi_k\}$,

$$\|T\|_{\mathrm{HS}}^2 := \sum_{j,k=1}^{\infty} \left|\langle \phi_j, T\phi_k\rangle\right|^2 < \infty. \tag{3.39}$$

(In the case of finite-dimensional matrices, this is called the Frobenius norm.) For integral operators, (3.38) shows that the sum on the right is independent of the choice of orthonormal basis. This remains true for any compact operator, although the proof is not so obvious.

3.6 Exercises

3.1. Let $T := \frac{d}{dx}$ acting on $L^2[0,1]$, as in Example 3.6. Compute $\mathcal{D}(T^*)$ for the following choices of domain:

(a) $\mathcal{D}(T) := C_0^\infty(0,1)$.
(b) $\mathcal{D}(T) := \{f \in C^\infty[0,1] : \ f(0) = f(1)\}$.

3.2. Consider the differentiable operator $T := -i\frac{d}{dx}$ for $x \in \mathbb{R}$.

(a) On $L^2(0,\infty)$, with the domain $\mathcal{D}(T) = C_0^\infty(0,\infty)$, show that T is symmetric but not essentially self-adjoint.
(b) On $L^2(\mathbb{R})$, with the domain $\mathcal{D}(T) = C_0^\infty(\mathbb{R})$, show that T is symmetric and essentially self-adjoint.

3.3. Prove Lemma 3.7: For an operator $T : \mathcal{D}(T) \to \mathcal{H}$,

$$\ker(T^*) = \operatorname{range}(T)^{\perp}.$$

3.4. Complete the proof of Theorem 3.18 by establishing the following: If the operator T is closed and bounded away from zero, then T is injective and range(T) is closed.

3.5. For an operator T on a Hilbert space $\mathcal{H}$, let $\mathcal{H}_T$ be the space $\mathcal{D}(T)$ equipped with the *graph norm*,

$$\|f\|_T := \left(\|f\|^2 + \|Tf\|^2\right)^{\frac{1}{2}}.$$

Prove that T is closed if and only if $\mathcal{H}_T$ is complete with respect to $\|\cdot\|_T$.

3.6. For a bounded self-adjoint operator A, prove that

$$\|A\| = \sup_{\|v\|=1} |\langle v, Av\rangle|.$$

[Hint: Use the easily verified identity,

$$\operatorname{Re}\langle v, Aw\rangle = \frac{1}{4}\Big[\big\langle v+w, A(v+w)\big\rangle - \big\langle v-w, A(v-w)\big\rangle\Big],$$

along with the result of Exercise 2.5.]

3.7. Suppose A is a positive operator. Prove that A is symmetric.

3.8. Prove that the closure of a symmetric operator is symmetric.

3.9. Prove that a closable operator A is essentially self-adjoint if and only if

$$\overline{A} = A^*.$$

3.10. Let T be a symmetric operator with self-adjoint extensions A_1 and A_2. If $\mathcal{D}(A_1) \subset \mathcal{D}(A_2)$, prove that $A_1 = A_2$.

3.11. Prove the following variant of Theorem 3.30: If A is a positive operator, then the following statements are equivalent:

(a) A is essentially self-adjoint.
(b) $A^* + 1$ is injective.
(c) $A + 1$ has dense range.

3.12. Let M_f be a multiplication operator on $L^2(X, d\mu)$, as in Example 3.2, with $f : X \to \mathbb{C}$ measurable and

$$\mathcal{D}(M_f) := \left\{ v \in L^2(X, d\mu) : \; fv \in L^2(X, d\mu) \right\}.$$

(X is assumed to be σ-finite.)

(a) Prove that the domain $\mathcal{D}(M_f)$ is dense in $L^2(X, d\mu)$.
(b) Prove that $\mathcal{D}(M_f^*) = \mathcal{D}(M_f)$ and $M_f^* = M_{\bar{f}}$. Hence M_f is self-adjoint if and only if f is real-valued a.e.

3.13. Given an orthonormal basis $\{\phi_k\}_{k=1}^{\infty}$ on a Hilbert space $\mathcal{H}$. let T be defined by

$$T\phi_k = \lambda_k \phi_k,$$

where $\{\lambda_k\}$ is a bounded sequence. Prove that T is compact if and only if $\lambda_k \to 0$.

3.14. Suppose that T is a compact operator on $\mathcal{H}$. Let $\{S_k\} \in \mathcal{L}(\mathcal{H})$ be a sequence of bounded operators such that $S_k^* \to 0$ in the strong operator sense. Prove that $TS_k \to 0$ in the operator norm topology.

Notes

Most of the standard texts on functional analysis or operator theory focus on the context of bounded operators on Banach spaces. We have chosen to emphasize unbounded operators from the beginning in order to develop the spectral theory of differential operators more quickly. This is similar to the approach taken in Weidmann [94], which gives a more complete treatment of the unbounded operator theory. For additional background on the unbounded case, see Riesz–Nagy [74, Chapter VIII], Kato [49, Chapter 5], Reed and Simon [69, Chapter VIII], or Schmüdgen [80].

Compact operators are a standard topic in any functional analysis text. See, for example, MacCluer [60, Chapter 4], for a presentation that includes the proof that the Hilbert–Schmidt norm (3.39) is independent of basis. For an introduction to compact operators that includes the theory of traces and determinants, see Simon [83, Chapter 3].

The relationship between operators and quadratic forms, which was used to define the Friedrichs extension in Section 3.4.3, is developed more systematically in Kato [49, Chapter 6], Reed and Simon [69, §VIII.6], Schmüdgen [80, Chapter 10], and Weidmann [94, §5.5].

There is a complete theory that describes the possible self-adjoint extensions of a symmetric operator, developed by von Neumann. We have chosen not to get into this here, because most of our later applications involve essentially self-adjoint operators. For details, see Reed and Simon [70, §X.1], Schmüdgen [80, Part VI], or Weidmann [94, Chapter 8].

Chapter 4
Spectrum and Resolvent



David Hilbert was the first to use the term "spectrum" to describe the set of eigenvalues of a linear operator, in a series of papers starting in 1904. He was apparently motivated by a loose analogy between the discrete sets of eigenvalues of certain integral operators and the atomic spectral lines discovered by physicists in the nineteenth century. Remarkably, Hilbert's analogy became a direct link just a few decades later, in Erwin Schrödinger's landmark calculation of the spectral lines of the hydrogen atom in 1926. (We will see this calculation in Section 7.4.)

The spectrum of a matrix M is the set of its eigenvalues (another term coined by Hilbert). Eigenvalues were originally defined by Augustin-Louis Cauchy as the roots of the characteristic polynomial, $q(z) := \det(M - zI)$. Since $q(\lambda) = 0$ precisely when $M - \lambda I$ fails to be invertible, each eigenvalue is associated with an eigenvector, a nonzero vector v for which $Mv = \lambda v$.

In the operator case, the determinant is not generally defined, so we define the spectrum in terms of the invertibility of $T - z$, for $z \in \mathbb{C}$. (To simplify the notation, a number appearing in an operator formula is interpreted as a multiple of the identity, so that $T - z$ stands for $T - zI$.) It still makes sense to define eigenvalues in terms of the existence of an eigenvector. This does not give the full spectrum, however, because there is no rank-nullity theorem in the infinite-dimensional case. The operator $T - \lambda$ could fail to be invertible even when $T - \lambda$ is injective and hence there is no eigenvector for λ.

4.1 Definitions and Examples

Let T be an operator on a Hilbert space $\mathcal{H}$ in the sense of Definition 3.1, i.e., possibly unbounded. An *eigenvalue* of T is a number $\lambda \in \mathbb{C}$ for which there exists a corresponding *eigenvector* $\phi \in \mathcal{D}(T)\backslash\{0\}$, such that

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1_4

$$T\phi = \lambda\phi.$$

The set of eigenvalues is called the *point spectrum* and denoted by $\sigma_{\mathrm{pt}}(T)$. Each eigenvalue carries a *multiplicity*, defined as the dimension of the space of eigenvectors sharing the same eigenvalue. In linear algebra terminology, this is the geometric multiplicity. (For the operator definition of algebraic multiplicity, see, for example, Gohberg and Krein [37, §1.1].)

As noted in the introduction, there might be more to the spectrum than eigenvalues. The full definition is as follows:

Definition 4.1. For an operator T, the *spectrum* $\sigma(T)$ is the set of points $\lambda \in \mathbb{C}$ for which $T - \lambda$ fails to have a bounded inverse. The complement of the spectrum is the *resolvent set*, denoted by $\rho(T)$. The bounded operator $(T - z)^{-1}$ is called the *resolvent* of T at $z \in \rho(T)$.

The "resolvent" is yet another term coined by Hilbert, but the concept had been introduced earlier by Ivar Fredholm in 1903. We will show in Section 4.2.2 that the resolvent set is open as a subset of $\mathbb{C}$, and therefore the spectrum is closed.

By the inverse mapping theorem (Theorem 3.17), the existence of a bounded inverse for $T - z$ requires T to be closed. Therefore, an operator that is not closed has all of $\mathbb{C}$ as its spectrum.

When T is closed, the inverse mapping theorem implies that $z \in \rho(T)$ if and only if $T - z$ is bijective. Thus there are only two possible ways for a point $\lambda \in \mathbb{C}$ to be contained in the spectrum. Either λ is an eigenvalue, or $T - \lambda$ is injective but not surjective.

Example 4.2. Consider the differential operator $T := -i\frac{d}{dx}$ from Example 3.21, which was seen to be self-adjoint on $L^2(0, 1)$ with the domain $\mathcal{D}(T)$ consisting of absolutely continuous functions satisfying periodic boundary conditions. This operator has the obvious eigenfunctions,

$$\phi_k(x) := e^{2\pi i k x}$$

for $k \in \mathbb{Z}$. The eigenfunctions constitute an orthonormal basis for $L^2(0, 1)$, by the Fourier basis arguments from Example 2.32.

For $z \in \mathbb{C}\backslash(2\pi\mathbb{Z})$, we can define the resolvent by

$$(T - z)^{-1} : \phi_k \mapsto (2\pi k - z)^{-1}\phi_k.$$

The operator $(T - z)^{-1}$ is thus bounded, with norm

$$\left\|(T - z)^{-1}\right\| = \frac{1}{\operatorname{dist}(z, 2\pi\mathbb{Z})}.$$

Hence, $\rho(T) = \mathbb{C}\backslash(2\pi\mathbb{Z})$ and

$$\sigma(T) = \sigma_{\mathrm{pt}}(T) = 2\pi\mathbb{Z}.$$

◊

4.1.1 Basic Properties of the Spectrum

For bounded operators we can use the operator norm to estimate the spectrum by working out a formula for $(T - z)^{-1}$ when $|z| > \|T\|$.

Lemma 4.3. *For a bounded operator T,*

$$\sigma(T) \subset \{z \in \mathbb{C} : |z| \leq \|T\|\}.$$

Proof For $|z| > \|T\|$, the series

$$S := \sum_{k=0}^{\infty} z^{-k} T^k \tag{4.1}$$

converges absolutely in $\mathcal{L}(\mathcal{H})$, and therefore defines a bounded operator. Clearly T commutes with S, and the obvious identity,

$$TS = z(S - I),$$

implies that

$$(T - z)S = -z.$$

Thus, $T - z$ has a bounded inverse and $z \in \rho(T)$. □

In later sections, we will study the resolvent operator in greater detail and use it to develop a better understanding of the spectrum. Before getting into this analysis, it is helpful to observe that the resolvent and spectrum behave naturally with respect to adjoints. Let us define the image of a subset $W \subset \mathbb{C}$ under complex conjugation by

$$W^* := \{z \in \mathbb{C} : \bar{z} \in W\}.$$

Theorem 4.4. *For a closed operator T,*

$$\sigma(T^*) = \sigma(T)^* \quad \textit{and} \quad \rho(T^*) = \rho(T)^*.$$

If $z \in \rho(T)$, then

$$\left[(T-z)^{-1}\right]^* = (T^* - \bar{z})^{-1}. \tag{4.2}$$

Proof It suffices to show that $\rho(T)^* \subset \rho(T^*)$. This will imply $\rho(T^*) \subset \rho(T)^*$ also, because $T^{**} = T$ for a closed operator.

Suppose $z \in \rho(T)$. For $v \in \mathcal{H}$ and $w \in \mathcal{D}(T^*)$,

$$\begin{aligned}\langle v, w\rangle &= \left\langle (T-z)(T-z)^{-1}v, w\right\rangle \\ &= \left\langle (T-z)^{-1}v, (T^*-\bar{z})w\right\rangle \\ &= \left\langle v, [(T-z)^{-1}]^*(T^*-\bar{z})w\right\rangle.\end{aligned}$$

Since v and w are arbitrary, this shows that

$$\left[(T-z)^{-1}\right]^*(T^*-\bar{z}) = I, \quad \text{on } \mathcal{D}(T^*).$$

A similar argument shows that

$$(T^*-\bar{z})\left[(T-z)^{-1}\right]^* = I, \quad \text{on } \mathcal{H}.$$

Hence $\bar{z} \in \rho(T^*)$ and (4.2) holds. □

4.1.2 Spectrum of a Multiplication Operator

The multiplication operators introduced in Example 3.2 will play a pivotal role in the spectral theorem developed in Chapter 5, analogous to the role played by diagonal matrices in linear algebra.

Let us recall how a multiplication operator is defined for a σ-finite measure space $(X, \mathcal{M}, \mu)$. Associated with a measurable function $f : X \to \mathbb{C}$ is an operator

$$M_f : u \mapsto fu,$$

acting on the Hilbert space $L^2(X, d\mu)$. This operator is bounded, with norm $\|M_f\| = \|f\|_\infty$, if and only if $f \in L^\infty(X, d\mu)$. Otherwise, M_f is defined as an unbounded operator on the domain

$$\mathcal{D}(M_f) := \left\{u \in L^2(X, d\mu) : fu \in u \in L^2(X, d\mu)\right\}.$$

The resolvent $(M_f - z)^{-1}$ at a point $z \in \mathbb{C}$ will clearly be given by the multiplication operator corresponding to $(f-z)^{-1}$, provided this is bounded. Therefore $z \in \rho(M_f)$ if and only if $(f-z)^{-1} \in L^\infty(X, d\mu)$. The spectrum of

M_f is thus closely related to the range of f. To describe it precisely, we introduce the notion of the *essential range* of f (with respect to μ), defined by

$$\text{ess-range}\, f := \Big\{ z \in \mathbb{C} : \ \mu(f^{-1}(B(z;\varepsilon))) > 0 \text{ for all } \varepsilon > 0 \Big\}.$$

This concept is quite similar to the essential supremum; both notions are unchanged when the function is modified on a set of measure zero. Note that for a continuous function on $\mathbb{R}^n$, the essential range (with respect to Lebesgue measure) is not the range itself, but rather its closure.

Theorem 4.5. *For a multiplication operator M_f on $L^2(X, d\mu)$,*

$$\sigma(M_f) = \text{ess-range}\, f.$$

If $z \in \rho(M_f)$, then

$$\left\| (M_f - z)^{-1} \right\| = \frac{1}{\operatorname{dist}(z, \sigma(M_f))}.$$

Furthermore, $\lambda \in \mathbb{C}$ is an eigenvalue of M_f if and only if $\mu(f^{-1}\{\lambda\}) > 0$.

Proof If $\lambda \in \text{ess-range}(F)$, then for all $\varepsilon > 0$, we have $|f - \lambda| < \varepsilon$ on a set of positive measure. This means that

$$\left\| (f - \lambda)^{-1} \right\|_\infty = \infty.$$

Thus $M_f - \lambda$ does not have a bounded inverse and hence $\lambda \in \sigma(M_f)$. This establishes

$$\text{ess-range}(f) \subset \sigma(M_f). \tag{4.3}$$

Now suppose that $z \notin \text{ess-range}(f)$. For $r < \operatorname{dist}(z, \text{ess-range}(f))$, this implies that

$$\mu(f^{-1}B(z;r)) = 0.$$

In other words, $|f(x) - z| \geq r$ for almost every $x \in X$, implying that

$$\left\| (f - z)^{-1} \right\|_\infty \leq r^{-1}, \tag{4.4}$$

and hence $z \in \rho(M_f)$. In view of (4.3), this proves

$$\text{ess-range}(f) \subset \sigma(M_f).$$

From (4.4) we also obtain the resolvent estimate,

$$\left\|(M_f - z)^{-1}\right\| \leq \operatorname{dist}(z, \sigma(M_f))^{-1}. \tag{4.5}$$

If $r > \operatorname{dist}(z, \text{ess-range}(f))$, then $|f - z| < r$ on a set of positive measure, and therefore

$$\left\|(f - z)^{-1}\right\|_{\infty} \geq r^{-1}.$$

Thus equality holds in (4.5).

Finally, let us prove the characterization of an eigenvalue. Suppose that $\phi \in L^2(X, d\mu)$ satisfies the eigenvalue equation,

$$(M_f - \lambda)\phi = 0. \tag{4.6}$$

In other words, $(f - \lambda)\phi = 0$ almost everywhere on X. If $\mu(f^{-1}\{\lambda\}) = 0$, then (4.6) implies that $\phi = 0$ almost everywhere. Therefore, a nontrivial solution of (4.6) is possible only if $\mu(f^{-1}\{\lambda\}) > 0$. Conversely, if $\mu(f^{-1}\{\lambda\}) > 0$, then there exists a set $E \subset f^{-1}\{\lambda\}$ with $0 < \mu(E) < \infty$. Setting $\phi = \chi_E$ yields an eigenfunction for λ. □

4.1.3 Resolvent of the Euclidean Laplacian

We saw in Example 3.23 that the Laplacian on $\mathbb{R}^n$ is self-adjoint with domain given by the Sobolev space $\mathcal{D}(-\Delta) = H^2(\mathbb{R}^n)$. Moreover, $-\Delta$ is unitarily equivalent to multiplication by $|\xi|^2$, under the Fourier transform. Therefore, by Theorem 4.5,

$$\sigma(-\Delta) = [0, \infty).$$

The unitary equivalence to $M_{|\xi|^2}$ also allows us to work out an explicit formula for the resolvent, which will prove useful later. For convenience, let us write the spectral parameter as $z = -\kappa^2$, so that $\operatorname{Re}\kappa > 0$ corresponds to the resolvent set. Our goal is to invert the operator $-\Delta + \kappa^2$.

By symmetry, the integral kernel of $(-\Delta + \kappa^2)^{-1}$ should depend only on the distance between points, so we seek a function $G(\kappa; r)$ such that

$$(-\Delta + \kappa^2)^{-1} f(x) = \int_{\mathbb{R}^n} G(\kappa; |x - y|) f(y)\, d^n y, \tag{4.7}$$

for $f \in L^2(\mathbb{R}^n)$. The integral kernel G is called a *Green's function* (or fundamental solution) for the operator $-\Delta + \kappa^2$.

To solve (4.7), we apply the Fourier transform to the expression on the left, yielding

$$\mathcal{F}\big[(-\Delta+\kappa^2)^{-1}f\big](\xi) = \frac{\hat{f}(\xi)}{|\xi|^2+\kappa^2}.$$

The inverse Fourier transform then gives

$$(-\Delta+\kappa^2)^{-1}f(x) = (2\pi)^{-\frac{n}{2}}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\frac{\hat{f}(\xi)}{|\xi|^2+\kappa^2}\,d^n\xi. \tag{4.8}$$

To compute the remaining integral, we write

$$\frac{1}{|\xi|^2+\kappa^2} = \int_0^\infty e^{-t(|\xi|^2+\kappa^2)}\,dt,$$

for $\operatorname{Re}\kappa^2 > 0$. Applying this to (4.8) and using the definition of $\hat{f}$, give (4.7) with

$$G(\kappa;|x-y|) := (2\pi)^{-n}\int_{\mathbb{R}^n}\int_0^\infty e^{i(x-y)\cdot\xi-t(|\xi|^2+\kappa^2)}\,d^n\xi\,dt$$

(where the change in integration order is justified by Fubini's theorem). The integral over ξ is now a standard Gaussian Fourier transform,

$$\int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi-t|\xi|^2}\,d\xi = \left(\frac{\pi}{t}\right)^{\frac{n}{2}} e^{-|x-y|^2/4t}.$$

Plugging this back into the expression for G gives

$$\begin{aligned} G(\kappa;r) &= (4\pi)^{-\frac{n}{2}}\int_0^\infty t^{-\frac{n}{2}}e^{-\kappa^2 t-r^2/4t}\,dt \\ &= (4\pi)^{-\frac{n}{2}}\kappa^{n-2}\int_0^\infty t^{-\frac{n}{2}}e^{-t-\kappa^2r^2/4t}\,dt. \end{aligned}$$

This expression for $G(\kappa;r)$ can be compared to a standard formula for the modified Bessel function [64, Eq. (10.32.10)],

$$K_\nu(w) := \frac{1}{2}\left(\frac{w}{2}\right)^\nu\int_0^\infty t^{-\nu-1}e^{-t-w^2/4t}\,dt,$$

valid for $\nu\in\mathbb{C}$ and $|\arg w| < \pi/4$. Setting $w=\kappa r$ gives

$$G(\kappa;r) = (2\pi)^{-\frac{n}{2}}\left(\frac{r}{\kappa}\right)^{1-\frac{n}{2}}K_{\frac{n}{2}-1}(\kappa r). \tag{4.9}$$

Although we assumed $\operatorname{Re}\kappa^2 > 0$ for sake of the calculation, we can now appeal to the Bessel equation to verify that the formula (4.9) extends to the full resolvent set $\operatorname{Re}\kappa > 0$. For $n = 3$, the formula (4.9) simplifies to

$$G(\kappa; r) = \frac{1}{4\pi}\frac{e^{-\kappa r}}{r}.$$

The classical Green's function is the integral kernel for $(-\Delta)^{-1}$. This can be recovered from (4.9) by formally taking $\kappa \to 0$. Using the well-known asymptotics of $K_\nu(w)$ as $w \to 0$, we obtain

$$G(0; r) = \begin{cases} -\dfrac{1}{2\pi}\log r, & n = 2, \\ \dfrac{1}{4}\pi^{-\frac{n}{2}}\Gamma(\frac{n}{2} - 1)r^{2-n}, & n \geq 3. \end{cases}$$

The existence of $G(0; r)$ as an integral kernel for $(-\Delta)^{-1}$ does not contradict the fact that $0 \in \sigma(-\Delta)$, because these kernels do not define bounded operators on $L^2(\mathbb{R}^n)$.

4.1.4 Discrete Laplacians

As a final class of examples, we consider some operators on the discrete Hilbert space $\ell^2(\mathbb{Z}^n)$ that arise in the context of solid-state physics.

To define a discrete analog of the Laplacian, we start by writing the one-dimensional Laplacian as the limit of a difference quotient,

$$-\Delta f(x) = \lim_{h\to 0} \frac{2f(x) - f(x+h) - f(x-h)}{h^2}.$$

This formula suggests that the discrete analog of the Laplacian on a regular lattice should be defined by freezing h at a nonzero value equal to the lattice spacing. For example, on $\ell^2(\mathbb{Z}^n)$ we would set $h = 1$ and define

$$-\Delta_{\mathbb{Z}^n} f(k) := \sum_{m\in\mathbb{Z}^n:\, |k-m|=1} (f(k) - f(m)). \tag{4.10}$$

Note that the sum is finite. For each $k \in \mathbb{Z}^n$, it includes the $2n$ nearest neighbors in the lattice. It is clear from the triangle inequality that

$$\|-\Delta_{\mathbb{Z}^n}\| \leq 4n.$$

We can therefore see that $-\Delta_{\mathbb{Z}^n}$ is self-adjoint just by checking symmetry,

$$\begin{aligned}\langle f, -\Delta_{\mathbb{Z}^n} g\rangle &= \sum_{k,m:|k-m|=1} \overline{f(k)}(g(k)-g(m)) \\ &= \sum_{k,m:|k-m|=1} \overline{(f(m)-f(k))}g(m) \\ &= \langle -\Delta_{\mathbb{Z}^n} f, g\rangle.\end{aligned}$$

To determine the spectrum of $-\Delta_{\mathbb{Z}^n}$, we use the discrete Fourier transform $\mathcal{F} : L^2(\mathbb{T}^n) \to \ell^2(\mathbb{Z}^n)$, given by

$$\mathcal{F}[h](k) = (2\pi)^{-n/2} \int_{\mathbb{T}^n} e^{-ik\cdot\theta} h(\theta)\, d^n\theta,$$

where $\mathbb{T}^n := (\mathbb{R}/2\pi\mathbb{Z})^n$. This map is unitary, with inverse given by the Fourier series,

$$\mathcal{F}^*[f](\theta) = (2\pi)^{-n/2} \sum_{k\in\mathbb{Z}^n} e^{ik\cdot\theta} f(k).$$

For $g \in L^2(\mathbb{T}^n)$, we compute

$$\begin{aligned}-\Delta_{\mathbb{Z}^n}\mathcal{F}[g](k) &= \sum_{|k-m|=1} (2\pi)^{-n/2} \int_{\mathbb{T}^n} \left(e^{-ik\cdot\theta} - e^{-im\cdot\theta}\right) g(\theta)\, d^n\theta \\ &= (2\pi)^{-n/2} \int_{\mathbb{T}^n} e^{-ik\cdot\theta} \sum_{j=1}^n \left(2 - e^{-i\theta_j} - e^{i\theta_j}\right) g(\theta)\, d^n\theta \\ &= \sum_{j=1}^n \mathcal{F}[(2-2\cos\theta_j)g](k).\end{aligned}$$

Thus $\mathcal{F}$ gives a unitary equivalence between $-\Delta_{\mathbb{Z}^n}$ and the multiplication operator M_H acting on $L^2(\mathbb{T}^n)$, for the function

$$H(\theta) := \sum_{j=1}^n (2 - 2\cos\theta_j).$$

It follows from Theorem 4.5 that

$$\sigma(-\Delta_{\mathbb{Z}^n}) = [0, 4n].$$

Discrete operators like $-\Delta_{\mathbb{Z}^n}$ are frequently used as models for problems in crystallography. These models are highly simplified from a physical point of view, and yet their mathematical behavior can be extremely complex. One of the more famous examples of this complexity arises as a discrete model for the problem of electrons in a two-dimensional lattice, subject to a constant magnetic field applied perpendicular to the lattice. Physicist P. G. Harper proposed in 1955 [42] to model this system using a quantum Hamiltonian operator H_α acting on $\ell^2(\mathbb{Z}^2)$, given by

$$H_\alpha u(m_1, m_2) := u(m_1 + 1, m_2) + u(m_1 - 1, m_2) \\ + e^{-2\pi i\alpha m_1} u(m_1, m_2 + 1) + e^{2\pi i\alpha m_1} u(m_1, m_2 - 1),$$

where $\alpha \in [0, 1]$. The parameter α represents the strength of the magnetic field. It is easy to check that H_α is bounded and self-adjoint. The $\alpha = 0$ case reduces to $H_0 = \Delta_{\mathbb{Z}^2} + 4$.

Harper's model is relatively easy to analyze when α is rational. If $\alpha = p/q$ with $p, q \in \mathbb{N}$, then H_α is periodic in both dimensions, with a fundamental cell of size $q \times 1$, as illustrated in Figure 4.1. To analyze the spectrum, we can adapt the discrete Fourier transform to this periodic structure. Define the map $U : \ell^2(\mathbb{Z}^2) \to L^2(\mathbb{T}^2, \mathbb{C}^q)$ by

$$(Uf)_j(\theta) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}^2} e^{-ik\cdot\theta} f(j - 1 + qk_1, k_2).$$

Note that the jth component of Uf is the standard Fourier transform of the restriction of f to the jth site in each cell. This map U is a special case of the *Floquet transform* (also called Bloch–Floquet), which is a fundamental tool in periodic spectral theory.

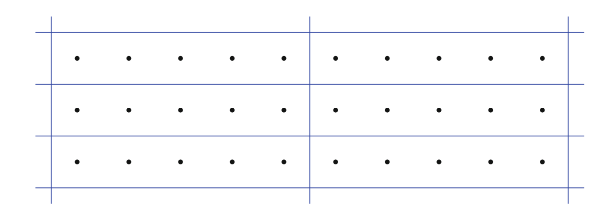


Fig. 4.1 Fundamental cell for H_α when $q = 5$

It is straightforward to check that U is unitary. To define its inverse, we write

$$U^* g(m_1, m_2) = \frac{1}{2\pi} \int_{\mathbb{T}^2} e^{ik\theta_1 + im_2\theta_2} g_j(\theta)\, d^2\theta,$$

where $m_1 = j - 1 + qk$, with $k \in \mathbb{Z}$ and $j = 0, 1, \ldots, q - 1$. The inversion formula for Fourier series implies that $U^* = U^{-1}$.

We can work out the spectrum of H_α (still assuming $\alpha = p/q$) by computing $UH_\alpha U^*$. It is helpful to write H_α in terms of shift operators, as

$$H_\alpha = S_1 + S_1^* + e^{-2\pi i p m_1/q} S_2 + e^{2\pi i p m_1/q} S_2^*, \tag{4.11}$$

where

$$S_1 u(m_1, m_2) := u(m_1 + 1, m_2), \qquad S_2 u(m_1, m_2) := u(m_1, m_2 + 1).$$

Under conjugation by U, S_2 behaves like an ordinary shift of Fourier series,

$$US_2U^* g(\theta) = e^{i\theta_2} g(\theta). \tag{4.12}$$

The action of S_1 is slightly more complicated. Setting $m_1 = j - 1 + qk$ as above, we have

$$\begin{aligned}(S_1U^*g)(\theta) &= U^*g(j + qk)\\ &= \begin{cases} \dfrac{1}{2\pi}\displaystyle\int_{\mathbb{T}^2} e^{ik\theta_1 + im_2\theta_2} g_{j+1}(\theta)\, d^2\theta, & j < q,\\ \dfrac{1}{2\pi}\displaystyle\int_{\mathbb{T}^2} e^{i(k+1)\theta_1 + im_2\theta_2} g_1(\theta)\, d^2\theta, & j = q.\end{cases}\end{aligned}$$

Applying U then gives

$$(US_1U^*g)j(\theta) = \begin{cases} g_{j+1}(\theta), & j < q,\\ e^{i\theta_1} g_1(\theta), & j = q.\end{cases} \tag{4.13}$$

By using (4.12) and (4.13) together with (4.11), we can see that U conjugates H_α to a matrix multiplication operator $\widetilde{A}$ on $L^2(\mathbb{T}^2, \mathbb{C}^q)$. This has the form

$$[\widetilde{A}g]_i(\theta) = \sum_{j=1}^{q} A_{ij}(\theta) g_j(\theta),$$

where $A(\theta)$ is the self-adjoint matrix-valued function

$$A(\theta) := \begin{pmatrix} 0 & e^{i\theta_1} \\ I_{q-1} & 0 \end{pmatrix} + \begin{pmatrix} 0 & I_{q-1} \\ e^{-i\theta_1} & 0 \end{pmatrix} + D(\theta_2),$$

with I_{q-1} the $(q-1)$-dimensional identity matrix, and $D(\theta_2)$ the $q \times q$ diagonal matrix given by

$$[D(\theta_2)]_{ij} = 2\cos\left(\frac{2\pi(j-1)p}{q} - \theta_2\right)\delta_{ij}.$$

Using the continuity of $A(\theta)$ and compactness of $\mathbb{T}^2$, it is straightforward to check that $\widetilde{A} - \lambda$ is invertible if and only if the 2×2 matrix $A(\theta) - \lambda$ is invertible for all $\theta \in \mathbb{T}^2$. Therefore, for $\alpha = p/q$ we may conclude that

$$\sigma(H_\alpha) = \bigcup_{\theta \in \mathbb{T}^2} \sigma(A(\theta)). \tag{4.14}$$

For example, in the case $p = 1$ and $q = 2$,

$$A(\theta) = \begin{pmatrix} 2\cos\theta_2 & 1 + e^{i\theta 1} \\ 1 + e^{-i\theta_2} & -2\cos\theta_2 \end{pmatrix}.$$

The eigenvalues of $A(\theta)$ are

$$\lambda_\pm(\theta) = \pm\sqrt{2 + 2\cos(\theta_1) + 4\cos^2(\theta_2)},$$

and therefore

$$\sigma(H_{1/2}) = \left[-\sqrt{8}, \sqrt{8}\right].$$

In his 1976 PhD thesis, Douglas Hofstadter [45] performed numerical calculations which first revealed the intricate structure of $\sigma(H_\alpha)$ for rational α. His work included a two-dimensional plot of these spectra, now known as the "Hofstadter butterfly," which is illustrated in Figure 4.2.

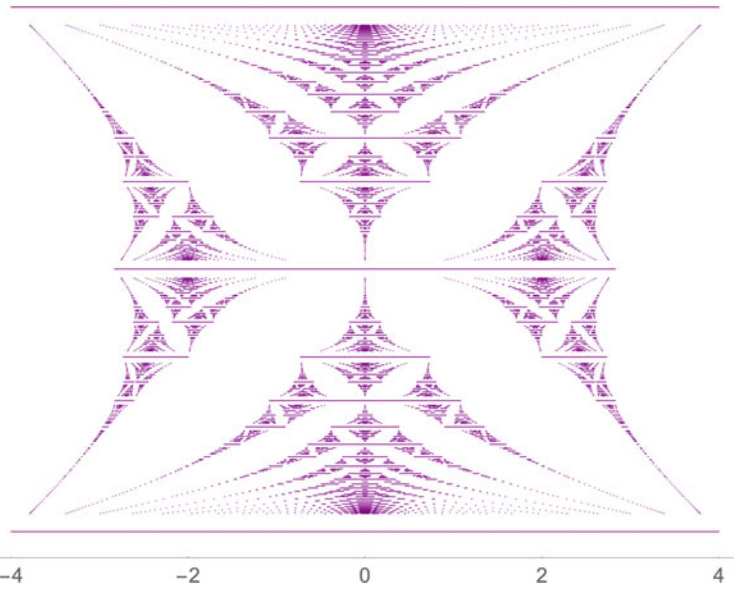


Fig. 4.2 The spectrum of H_α is plotted horizontally for rational values of α ranging from 0 to 1 on the vertical axis

This plot illustrates the possibility of extreme instability of the spectrum under perturbations. One might imagine that the intricate structure of the spectral bands is an artifact of Harper's simplified model. However, in 2013 such patterns were observed in nature by several independent research groups. These experiments involved graphene, an exotic material consisting of carbon atoms arranged in a two-dimensional hexagonal lattice that is only one atom thick.

Determining the spectrum of H_α for α irrational is a much more difficult problem than the calculation used to create Figure 4.2. Hofstadter's work gave support to an earlier conjecture by Mark Azbel that $\sigma(H_\alpha)$ is a Cantor set for irrational values of α, and thus has measure zero. This conjecture, later popularized by Mark Kac and Barry Simon as the "ten-martini" problem, was finally proven by Artur Avila and Svetlana Jitomirskaya in 2009 [6].

4.2 Resolvent

For a matrix, the fact that the spectrum can be defined as the roots of the characteristic polynomial allows us to apply powerful tools from complex analysis to spectral questions. For example, the existence of eigenvalues follows from Liouville's theorem, via the fundamental theorem of algebra.

For a general operator there is no direct analog of the characteristic polynomial. However, we can still bring complex analysis into the picture by interpreting the resolvent as a holomorphic function of the spectral parameter.

4.2.1 Analytic Operator-Valued Functions

To set up our analysis of the resolvent, we first develop some basic theory of analytic families of bounded operators. Throughout this discussion, we will use Ω to denote an open, connected subset of $\mathbb{C}$.

Definition 4.6. A map $F : \Omega \to \mathcal{L}(\mathcal{H})$ is *analytic* if for each $z_0 \in \Omega$ there exists a sequence of bounded operators $\{A_n\}_{n=1}^{\infty} \subset \mathcal{L}(\mathcal{H})$ and $r_0 > 0$ such that

$$F(z) = \sum_{n=1}^{\infty} (z - z_0)^n A_n, \tag{4.15}$$

with the series converging absolutely for $|z - z_0| < r_0$.

By the root test, the radius of absolute convergence of the power series (4.15) is

$$r_0 = \left(\limsup_{n\to\infty} \|A_n\|^{\frac{1}{n}} \right)^{-1}.$$

In particular this means that for $T \in \mathcal{L}(\mathcal{H})$, the geometric series,

$$(I - zT)^{-1} = \sum_{n=0}^{\infty} z^n T^n, \tag{4.16}$$

is convergent for $|z| < \|T\|^{-1}$. An operator expansion of the form (4.16) is called a *Neumann series*. It is easy to check that $(I - zT)^{-1}$ is analytic at each point $z_0 \in \{|z| < \|T\|^{-1}\}$, using a geometric series expansion:

$$\begin{aligned}(I - zT)^{-1} &= [(I - z_0 T) - (z - z_0)T]^{-1} \\ &= (I - z_0 T)^{-1} \sum_{k=0}^{\infty} (z - z_0)^k \big[T(I - z_0 T)^{-1}\big]^k,\end{aligned}$$

for $|z - z_0|$ sufficiently small.

As in the scalar case, operator-valued analyticity is related to the existence of a complex derivative. A function $F : \Omega \to \mathcal{L}(\mathcal{H})$ is *holomorphic* if the limit

$$F'(z) := \lim_{h \to 0} \frac{F(z+h) - F(z)}{h}, \tag{4.17}$$

exists (in the operator-norm topology) for each $z \in \Omega$. To make the connection between holomorphic and analytic functions, we can use operator-valued contour integrals.

To keep this discussion brief, we will only define contour integrals in a weak sense. We say that $F : \Omega \to \mathcal{L}(\mathcal{H})$ is *weakly continuous* if the function $z \mapsto \langle v, F(z)w \rangle$ is continuous for each $v, w \in \mathcal{H}$.

Lemma 4.7. *Suppose that $F : \Omega \to \mathcal{L}(\mathcal{H})$ is weakly continuous, and let γ be a closed, piecewise smooth curve in Ω. There is a unique bounded operator, denoted by*

$$T := \int_\gamma F(z)\, dz \tag{4.18}$$

such that

$$\langle v, Tw \rangle = \int_\gamma \langle v, F(z)w \rangle \, dz$$

for all $v, w \in \mathcal{H}$. Moreover,

$$\|T\| \leq \ell(\gamma) \sup_{z \in \gamma} \|F(z)\|.$$

Proof The sesquilinear form,

$$q(v,w) := \int_\gamma \langle v, F(z)w\rangle \, dz,$$

is well defined by the weak continuity assumption. Furthermore, since γ is compact,

$$\sup_{z\in\gamma} \langle v, F(z)w\rangle < \infty$$

for each $v, w \in \mathcal{H}$. The uniform boundedness principle (Theorem 2.13) implies that

$$\sup_{z\in\gamma} \|F(z)\| < \infty.$$

Hence $q(\cdot,\cdot)$ is bounded as a sesquilinear form, with

$$\|q(\cdot,\cdot)\| \le \ell(\gamma) \sup_{z\in\gamma} \|F(z)\|. \tag{4.19}$$

By Corollary 2.29 of the Riesz lemma, it follows that there exists a unique $T \in \mathcal{L}(\mathcal{H})$ such that

$$q(v,w) = \langle v, Tw\rangle,$$

with $\|T\| = \|q(\cdot,\cdot)\|$. □

One could extend the definition of (4.18) through an operator-valued formulation of Riemann sums. This approach allows for considerably less restrictive regularity assumptions on F and γ. We will not go into the details here, as the weak formulation given in Lemma 4.7 is sufficient for our applications.

We say that a function $F : \Omega \to \mathcal{L}(\mathcal{H})$ is *weakly holomorphic* if, for each $v, w \in \mathcal{H}$, the function

$$z \mapsto \langle v, F(z)w\rangle$$

is holomorphic on Ω. The advantage in using weak notions of holomorphicity and contour integration is that we can develop the theory using ordinary complex analysis. The following result shows that we do not lose anything by taking this approach.

Theorem 4.8. *For a function $F : \Omega \to \mathcal{L}(\mathcal{H})$, these conditions are equivalent:*

(a) *F is analytic.*
(b) *F is holomorphic.*
(c) *F is weakly holomorphic.*

Proof To prove that (a) implies (b), one must show that term-by-term differentiation of the power series (4.15) is justified. This argument is essentially the same as in the scalar case, so we omit the details.

The fact that (b) implies (c) is trivial, so it remains to prove that (c) implies (a). Assume that F is weakly holomorphic on Ω. Let $z_0 \in \Omega$. Choose $r > 0$ so that $\overline{B(z_0; r)} \subset \Omega$, and let $\gamma := \partial B(z_0; r)$. Applying Cauchy's formula to the scalar holomorphic function $\langle v, F(\cdot)w\rangle$ gives an expansion

$$\langle v, F(z)w\rangle = \sum_{n=0}^{\infty} (z - z_0)^n a_n(v, w), \tag{4.20}$$

where

$$a_n(v, w) := \frac{1}{2\pi i} \int_\gamma \frac{\langle v, F(z)w\rangle}{(z - z_0)^{n+1}} dz \tag{4.21}$$

for $n \in \mathbb{Z}_0$.

By Lemma 4.7, the coefficient (4.21) can be written as

$$a_n(v, w) = \langle v, A_n w\rangle,$$

where

$$A_n := \frac{1}{2\pi i} \int_\gamma \frac{F(z)}{(z - z_0)^{n+1}} dz.$$

Lemma 4.7 also gives the estimate,

$$\|A_n\| \le \frac{1}{r^n} \sup_{z \in \gamma} \|F(z)\|,$$

implying that the power series

$$Q(z) := \sum_{n=0}^{\infty} (z - z_0)^n A_n$$

converges absolutely for $|z - z_0| < r$. By (4.20),

$$\langle v, F(z)w\rangle = \langle v, Q(z)w\rangle$$

for all $v, w \in \mathcal{H}$. Hence $F(z) = Q(z)$ for $|z - z_0| < r$, and so $F(z)$ is analytic at z_0. □

As a corollary of the proof of Theorem 4.8, we obtain an operator form of Cauchy's derivative estimate.

Theorem 4.9. *Suppose $F : \Omega \to \mathcal{L}(\mathcal{H})$ is analytic, and $\overline{B(z_0; r)} \subset \Omega$. Then the derivatives of F at z_0 satisfy*

$$\left\| F^{(n)}(z_0) \right\| \le \frac{M_r n!}{r^n}$$

for $n \in \mathbb{N}$, where

$$M_r := \sup_{|z-z_0|=r} \|F(z)\|.$$

4.2.2 Analyticity of the Resolvent

We can now apply the theory introduced in Section 4.2.1 to the resolvent, and use it to derive basic properties of the spectrum.

Theorem 4.10. *For a closed operator T, the resolvent set $\rho(T)$ is open and the function $z \mapsto (T - z)^{-1}$ is analytic on each connected component of $\rho(T)$.*

Proof For convenience in this discussion, let us write the resolvent as

$$R(z; T) := (T - z)^{-1}. \tag{4.22}$$

To see that $T - z$ is invertible for z sufficiently close to an arbitrary point $z_0 \in \rho(T)$, we first note that

$$(T - z)R(z_0; T) = I - (z - z_0)R(z_0; T). \tag{4.23}$$

The right-hand side can be inverted using the Neumann series,

$$Q(z) := \sum_{n=0} (z - z_0)^n R(z_0; T)^n,$$

which defines an analytic function for $|z - z_0| < \|R(z_0; T)\|^{-1}$. It follows from (4.23) that

$$(T - z)Q(z)R(z_0; T) = I. \tag{4.24}$$

Similarly, on the domain $\mathcal{D}(T)$ we have

$$R(z_0; T)(T - z) = I - (z - z_0)R(z_0; T),$$

Applying $Q(z)$ to this relation gives

$$Q(z)R(z_0; T)(T - z) = I \tag{4.25}$$

on $\mathcal{D}(T)$.

Together, (4.24) and (4.25) imply that

$$R(z;T) = Q(z)R(z_0;T).$$

By the definition of $Q(z)$, this yields an expansion

$$R(z;T) = \sum_{n=0}^{\infty} (z - z_0)^n R(z_0;T)^{n+1} \tag{4.26}$$

for $|z - z_0| < \|R(z_0;T)\|^{-1}$. Thus $R(\cdot;T)$ is analytic at z_0. □

Theorem 4.10 has several important corollaries. The first is a simple observation of the relationship between the norm of the resolvent and the radius of convergence in the expansion (4.26).

Corollary 4.11. *For $z \in \rho(T)$,*

$$\left\|(T - z)^{-1}\right\| \geq \frac{1}{\operatorname{dist}(z, \sigma(T))}. \tag{4.27}$$

The inequality (4.27) proves to be quite useful in both directions. When the spectrum is known it provides an easy bound on the resolvent. On the other hand, if the spectrum is unknown, then estimates of the resolvent can be used to help locate it. We will see in Section 5.3 that (4.27) becomes an equality in the case of self-adjoint operators.

The expansion formula (4.26) also yields a relation that makes it easy to compare the resolvent evaluated at different values if the argument.

Corollary 4.12 (First Resolvent Identity). *For an operator $T : \mathcal{D}(T) \to \mathcal{H}$, the resolvents at points $z, w \in \rho(T)$ commute and satisfy*

$$(T - z)^{-1} - (T - w)^{-1} = (z - w)(T - z)^{-1}(T - w)^{-1}.$$

There is also a second resolvent identity: if $\mathcal{D}(S) = \mathcal{D}(T)$ and $z \in \rho(S) \cap \rho(T)$, then

$$(S - z)^{-1} - (T - z)^{-1} = (S - z)^{-1}(S - T)(T - z)^{-1}.$$

See Exercise 4.1 for the proof.

4.2.3 Spectral Radius

Another application of Theorem 4.10 concerns the *spectral radius* of an operator T, defined by

$$r(T) := \sup_{z \in \sigma(T)} |z|.$$

For a self-adjoint matrix the spectral radius is equal to the norm, and we would expect these quantities to be closely related for operators as well.

The spectral radius of a bounded operator can be estimated using the analog of a Laurent series expansion for the resolvent. For $T \in \mathcal{L}(\mathcal{H})$, replacing z by $1/z$ in the Neumann series formula (4.16) gives

$$(T - z)^{-1} = \sum_{n=0}^{\infty} z^{-n-1} T^n, \tag{4.28}$$

for $|z| > \|T\|$.

Corollary 4.13. *For a bounded operator T, the spectrum is not empty and the spectral radius satisfies*

$$r(T) \leq \|T\|.$$

Proof The bound on the spectral radius follows immediately from the expansion (4.28). To prove the first claim, suppose that $\sigma(T)$ is empty. This implies that $(T - z)^{-1}$ is analytic for all $z \in \mathbb{C}$ by Theorem 4.10. For $v, w \in \mathcal{H}$, the function,

$$h(z) := \left\langle (T - z)^{-1} v, w \right\rangle,$$

is then entire and satisfies $h(z) \to 0$ as $z \to \infty$ by (4.28). Therefore $h \equiv 0$ by Liouville's theorem. Since this holds for any v, w, it implies that $(T - z)^{-1} = 0$ for all z, which is not possible. □

The relationship between the norm of a bounded operator and its spectral radius was made more precise by I. M. Gelfand [35].

Theorem 4.14 (Gelfand Spectral Radius Formula). *For a bounded operator T,*

$$r(T) = \lim_{n \to \infty} \|T^n\|^{\frac{1}{n}}.$$

Furthermore, if T is bounded and self-adjoint then $r(T) = \|T\|$.

Proof If $T^n - z^n$ is invertible for $z \in \mathbb{C}$ and $n \in \mathbb{N}$, then $T - z$ is also invertible, by

$$(T - z)^{-1} = (T^n - z^n)^{-1} (T^{n-1} + z T^{n-2} + \cdots + z^{n-1}).$$

Therefore, if $\lambda \in \sigma(T)$, then $\lambda^n \in \sigma(T^n)$, and we have

$$|\lambda|^n \leq \|T^n\|.$$

Taking the nth root and letting $n \to \infty$ show that

$$r(T) \le \liminf_{n\to\infty} \|T^n\|^{\frac{1}{n}}. \tag{4.29}$$

On the other hand, consider the function $F(z) := (I - zT)^{-1}$. This is analytic for near $z = 0$ by the Neumann series expansion (4.16), and for $0 < |z| < r(T)^{-1}$ by the identity

$$F(z) = -z(T - 1/z)^{-1}.$$

Applying Theorem 4.9 to the disk $\{|z| \le a\}$ for $a < r(T)^{-1}$ gives the estimate

$$\|T^n\| \le M_\varepsilon a^{-n}$$

for $n \in \mathbb{N}$. In the limit $n \to \infty$ this gives

$$\limsup_{n\to\infty} \|T^n\|^{\frac{1}{n}} \le a^{-1},$$

for all $a < r(T)^{-1}$. Hence,

$$\limsup_{n\to\infty} \|T^n\|^{\frac{1}{n}} \le r(T).$$

In combination with (4.29) this proves that $\|T^n\|^{\frac{1}{n}} \to r(T)$ as $n \to \infty$.

If $A \in \mathcal{L}(\mathcal{H})$ is self-adjoint, then $\|A^2\| = \|A\|^2$ by (3.4). Iterating this result, we see that

$$\left\|A^{2^k}\right\| = \|A\|^{2^k}$$

for $k \in \mathbb{N}$. Restricting the limit in the spectral radius formula to the subsequence $n = 2^k$ then gives $r(A) = \|A\|$. □

For general unbounded operators, there is no analog of Corollary 4.13 or Theorem 4.14. The spectrum of an unbounded operator could be empty, for example. However, for a self-adjoint operator A we will see in Chapter 5 that $\sigma(A)$ is not empty and that $r(A) = \infty$ if and only if A is unbounded (see Theorem 5.9).

4.3 Spectrum of Self-adjoint Operators

Certain basic properties of the spectrum of a self-adjoint matrix carry over directly to the operator case.

Theorem 4.15. *The spectrum of a self-adjoint operator is real, and eigenvectors corresponding to distinct eigenvalues are orthogonal.*

Proof If A is closed, then $A - z$ has a bounded inverse if and only if $A - z$ is bijective, by Theorem 3.17. Thus, for A self-adjoint and z strictly complex, Theorem 3.29 implies that $z \in \rho(A)$. Hence $\sigma(A) \subset \mathbb{R}$.

Now suppose that ϕ_1 and ϕ_2 are eigenvectors of A, with corresponding eigenvalues λ_1, λ_2. Because the eigenvalues are real,

$$\begin{aligned} 0 &= \langle A\phi_1, \phi_2 \rangle - \langle \phi_1, A\phi_2 \rangle \\ &= (\lambda_1 - \lambda_2)\langle \phi_1, \phi_2 \rangle. \end{aligned}$$

Hence $\lambda_1 \neq \lambda_2$ implies that ϕ_1 is orthogonal to ϕ_2. □

Another important feature of self-adjoint operators is the fact that each point in the spectrum is an "approximate eigenvalue" in the following sense.

Theorem 4.16. *Let A be a self-adjoint operator. Then $z \in \sigma(A)$ if and only if there exists a sequence $\{u_n\} \subset \mathcal{D}(A)$ with $\|u_n\| = 1$ and*

$$\lim_{n \to \infty} \|(A - z)u_n\| = 0.$$

Proof If $z \in \rho(A)$, then $(A - z)^{-1}$ maps $\mathcal{H}$ onto $\mathcal{D}(A)$. For $u \in \mathcal{D}(A)$ we can thus set $u = (A - z)^{-1}v$ for some $v \in \mathcal{H}$. If $\|u\| = 1$, then the boundedness of $(A - z)^{-1}$ gives

$$1 \le \left\| (A - z)^{-1} \right\| \|v\|.$$

Hence

$$\|(A - z)u\| \ge \frac{1}{\|(A - z)^{-1}\|}$$

for all unit vectors $u \in \mathcal{D}(T)$. Thus there cannot exist a sequence of unit vectors for which $(A - z)u_n \to 0$.

Now assume $\lambda \in \sigma(A)$. If λ is an eigenvalue, then we can trivially set each u_n equal to an eigenvector. So let us assume that λ is not an eigenvalue, meaning that $A - \lambda$ is injective but not surjective. In this case, by Lemma 3.7 and the fact that A is self-adjoint,

$$\operatorname{range}(A - \lambda)^{\perp} = \{0\}.$$

Hence $\operatorname{range}(A - \lambda)$ is dense in $\mathcal{H}$.

Since $A - \lambda$ is injective, there exists a linear map $W : \operatorname{range}(A - \lambda) \to \mathcal{D}(A)$ such that $(A - \lambda)W = I$ on $\operatorname{range}(A - \lambda)$. Because $\operatorname{range}(A - \lambda)$ is dense, the operator W is unbounded. Otherwise W could be extended to $\mathcal{H}$ by continuity, which would imply $\lambda \in \rho(A)$. Because W is unbounded there exists a sequence $\{v_n\} \subset \operatorname{range}(A - \lambda)$ such that $\|v_n\| = 1$ and

$$\lim_{n \to \infty} \|W v_n\| = \infty. \tag{4.30}$$

From the sequence $\{v_n\}$ we construct a sequence of unit vectors,

$$u_n := \frac{Wv_n}{\|Wv_n\|},$$

satisfying

$$\|(A-\lambda)u_n\| = \frac{1}{\|Wv_n\|}.$$

Then, by (4.30),

$$\lim_{n\to\infty} \|(A-\lambda)u_n\| = 0.$$

□

Results such as Theorem 4.16 are particularly useful for perturbations of an operator whose spectrum is already known. This situation is common in physical applications.

Example 4.17. In quantum mechanics, a *Schrödinger operator* on $\mathbb{R}^n$ is a differential operator of the form $-\Delta + V$ where V is a potential function acting by multiplication. Suppose that V is a bounded, real-valued function with compact support. Then multiplication by V defines a bounded operator, and hence $-\Delta + V$ is self-adjoint on the same domain as $-\Delta$, namely $H^2(\mathbb{R}^n)$.

We can construct approximating sequences for $\sigma(-\Delta + V)$ from plane wave functions $e^{i\xi\cdot x}$, for $\xi \in \mathbb{R}^n$, which satisfy

$$-\Delta e^{i\xi\cdot x} = |\xi|^2 e^{i\xi\cdot x}.$$

Choose $\psi \in C^\infty(\mathbb{R})$ with $\psi(t) = 1$ for $t \le 0$ and $\psi(t) = 0$ for $t \ge 1$. Then define the family of cutoffs $\chi_k \in C_0^\infty(\mathbb{R}^n)$ by

$$\chi_k(x) := \psi(|x| - k),$$

for $k \in \mathbb{N}$. For the sequence

$$f_k(x) := \chi_k(x) e^{i\xi\cdot x},$$

we have

$$(-\Delta + V - |\xi|^2) f_k = V f_k - [\Delta, \chi_k] e^{i\xi\cdot x},$$

where the commutator,

$$[\Delta, \chi_k] := \Delta\chi_k - \chi_k\Delta,$$

is a first-order differential operator. For k large, $|V f_k| = |V|$, and we can estimate

$$\begin{aligned}\left\|[\Delta, \chi_k]e^{i\xi\cdot x}\right\|^2 &\le C_\xi \operatorname{vol}\{k \le |x| \le k+1\}\\ &= O(k^{n-1}).\end{aligned}$$

Thus

$$\left\|(-\Delta + V - |\xi|^2) f_k\right\| = O(k^{(n-1)/2}).$$

Since $\|f_k\|^2 \ge \operatorname{vol} B(0; k)$, we also have

$$\|f_k\| \ge ck^{n/2}.$$

Therefore $f_k/\|f_k\|$ defines an approximate eigenfunction sequence for $-\Delta + V$, with eigenvalue $|\xi|^2$). Theorem 4.16 implies that

$$[0, \infty) \subset \sigma(-\Delta + V).$$

We will refine this argument in Section 5.4, and Schrödinger operators are studied in greater detail in Chapter 7. $\Diamond$

4.4 Spectral Theory of Compact Operators

We saw in Theorem 3.37 that compact operators are operator-norm limits of finite-rank operators. It is therefore to be expected that the spectral theory of compact operators will resemble the matrix theory. In this section we establish a number of spectral properties for compact operators which are clear analogs of linear algebra results.

The first step is to show that eigenspaces of a compact operator are finite-dimensional.

Lemma 4.18. *If $T \in \mathcal{L}(\mathcal{H})$ is compact and $\lambda \neq 0$, then*

$$\dim \ker(T - \lambda) < \infty.$$

Proof Suppose that $\ker(T - \lambda)$ is infinite-dimensional. Then there exists an orthogonal sequence $\{e_j\}_{j=1}^\infty$ such that $Te_j = \lambda e_j$. If $\lambda \neq 0$ the sequence $\{\lambda e_j\}$ clearly has no convergent subsequence, so T is not compact. $\square$

To analyze the resolvent $(T - z)^{-1}$ for $z \neq 0$, it is equivalent to study the invertibility of $I - zT$ for $z \in \mathbb{C}$. This is the immediate application that we have in mind for the following result. However, the more general formulation, in terms of analytic families of compact operators, proves to be quite useful.

Theorem 4.19 (Analytic Fredholm Theorem). *Let $\mathcal{H}$ be a separable Hilbert space, and $F : \Omega \to \mathcal{L}(\mathcal{H})$ an analytic function on an open connected domain $\Omega \subset \mathbb{C}$, such that $F(z)$ is compact for each $z \in \Omega$. Then either:*

(a) *$I - F(z)$ fails to be invertible for every $z \in \Omega$; or*
(b) *there exists a discrete subset $\Gamma \subset \Omega$ such that $I - F(z)$ is invertible for $z \in \Omega \backslash \Gamma$ and $\ker[I - F(z)] \neq \{0\}$ for $z \in \Gamma$.*

Proof Let z_0 be a point in Ω. Our first goal is to reduce the invertibility of $I - F(z)$ to a finite-dimensional problem near z_0. By the continuity of $F(\cdot)$ there exists some $\varepsilon > 0$ such that

$$\|F(z) - F(z_0)\| < \frac{1}{2}$$

for $|z - z_0| < \varepsilon$. By Theorem 3.37, there exists a finite-rank operator R such that

$$\|F(z_0) - R\| < \frac{1}{2}.$$

Thus, by the triangle inequality,

$$\|F(z) - R\| < 1$$

for $z \in B(z_0; \varepsilon)$.

The operator

$$Q(z) := I - F(z) + R$$

is therefore invertible by Neumann series for $z \in B(z_0; \varepsilon)$. Furthermore, the fact that $F(z)$ is analytic implies that $Q(z)^{-1}$ is analytic. We can see this by deriving the power series coefficients for $Q(z)^{-1}$ from those of $Q(z)$, just as in the scalar case.

The invertibility of $Q(z)$ can be exploited to write

$$\begin{aligned} I - F(z) &= Q(z) - R \\ &= \left(I - RQ(z)^{-1}\right)Q(z). \end{aligned} \tag{4.31}$$

Hence $I - F(z)$ is invertible if and only if $I - RQ(z)^{-1}$ is invertible. Moreover, $Q(z)$ maps $\ker[I - F(z)]$ bijectively onto $\ker[I - RQ(z)^{-1}]$.

Let $\{e_1, \dots, e_n\}$ be an orthonormal basis for range(R), so that

$$Rv = \sum_{j=1}^{n} \langle e_j, Rv\rangle e_j.$$

Define the matrix-valued analytic function $M(z)$ by

$$[M(z)]_{ij} := \left\langle e_i, RQ(z)^{-1} e_j \right\rangle.$$

The determinant

$$D(z) := \det(I - M(z))$$

is analytic for $z \in B(z_0; \varepsilon)$. Hence either $D \equiv 0$ or D has a discrete set of zeros in this disk.

The next step is to show that $D(z) = 0$ precisely when $I - F(z)$ fails to be invertible. If $D(z) = 0$, then there exists a nonzero vector $(a_1, \dots, a_n) \in \mathbb{C}^n$ such that

$$a_i = \sum_{j=1}^{n} \langle e_i, RQ(z)^{-1} e_j \rangle a_j. \tag{4.32}$$

For $v := \sum_{j=1}^{n} a_j e_j$ this is equivalent to

$$v = RQ(z)^{-1} v.$$

By (4.31),

$$(I - F(z))Q(z)v = 0.$$

Hence $D(z) = 0$ implies $\ker[I - F(z)] \neq \{0\}$. If, on the other hand, $D(z) \neq 0$, then (4.32) has no nontrivial solutions. Hence $I - RQ(z)^{-1}$ is invertible and so is $I - F(z)$.

At this stage, we have shown that the claimed result holds in a neighborhood of each point $z_0 \in \Omega$. The proof is completed with a standard connectivity argument. The sets

$$A := \left\{ z \in \Omega : \ I - F(\cdot) \text{ is not invertible in a neighborhood of } z \right\}$$

and

$$B := \left\{ z \in \Omega : I - F(\cdot) \text{ is invertible in a neighborhood of } z \text{ minus a discrete set} \right\}$$

are open and disjoint by definition, and we have shown that $\Omega = A \cup B$. Since Ω is connected, this implies that either $\Omega = A$ or $\Omega = B$. □

As a corollary of the analytic Fredholm theorem, we obtain the following characterization of the spectrum of a compact operator.

Theorem 4.20 (Riesz–Schauder Theorem). *For a compact operator T on a separable Hilbert space $\mathcal{H}$, all elements of $\sigma(T)\backslash\{0\}$ are eigenvalues of finite multiplicity, and $\sigma(T)$ has no limit point other than possibly* 0.

Proof Clearly $I - zT$ is invertible at $z = 0$, so applying Theorem 4.19 with $F(z) = zT$ shows that there exists a discrete subset $\Gamma \subset \mathbb{C}$ such that $I - zT$ is invertible on $\mathbb{C}\backslash\Gamma$ and $\ker(1 - zT) \neq \{0\}$ for z in Γ. From the relation

$$I - zT = -z(T - z^{-1}),$$

it follows that the points of $\sigma(T)\backslash\{0\}$ are the reciprocals of the points in Γ, and that each of these is an eigenvalue. The multiplicities of these eigenvalues are finite by Lemma 4.18. □

4.4.1 Spectral Theorem for Compact Self-adjoint Operators

Theorem 4.20 does not guarantee the existence of eigenvalues. Although we know that the spectrum of a compact operator is nonempty, by Corollary 4.13, it is possible that the spectrum equals $\{0\}$ and that zero is not an eigenvalue. For example, given an orthonormal basis $\{e_k\}$, consider the modified shift operator

$$Te_k = \frac{1}{k} e_{k+1}.$$

If the compact operator is self-adjoint, however, its spectral theory closely resembles the matrix case.

Theorem 4.21 (Hilbert–Schmidt theorem). *Let A be a compact self-adjoint operator on a separable Hilbert space $\mathcal{H}$. There exists an orthonormal basis $\{\phi_k\}$ for $\mathcal{H}$, such that*

$$A\phi_k = \lambda_k \phi_k$$

for $\lambda_k \in \mathbb{R}$, with $\lambda_k \to 0$ as $k \to \infty$.

Proof For each eigenvalue of A, choose an orthonormal basis for the corresponding eigenspace. By Theorem 4.15 and the separability of $\mathcal{H}$, the collection of these basis elements forms a (possibly finite) orthonormal sequence $\{\phi_k\}$. If there are infinitely many eigenfunctions, then the fact that the eigenvalues converge to zero follows from Theorem 4.20.

Let $\mathcal{W}$ be the closure of the span of $\{\phi_n\}$. It is easy to check that since A preserves W and is self-adjoint, it also preserves $\mathcal{W}^\perp$. The restriction of A to $\mathcal{W}^\perp$ thus defines a new operator $A_\perp$, which is also compact and self-adjoint. Moreover, $A_\perp$ has no eigenvectors, because the eigenvectors of A were all included in $\mathcal{W}$. By Theorem 4.20 we conclude that $A_\perp$ has spectral radius $r(A_\perp) = 0$, and it then follows from Theorem 4.14 that $A_\perp = 0$. This implies also that $\mathcal{W}^\perp = \{0\}$, because any nonzero element of $\mathcal{W}^\perp$ would be an eigenvector of A with eigenvalue 0. Therefore $\mathcal{W} = \mathcal{H}$, implying that $\{\phi_n\}$ is a basis. □

Theorem 4.21 is a special case of the more general spectral theorem to be proven in Chapter 5. For computations or estimation of the eigenvalues of a compact

operator, we can adapt yet another basic principle from the matrix case. (See Section 5.4.3 for a more general operator version.)

Theorem 4.22 (Max–Min Principle for Compact Operators). *Suppose that A is a self-adjoint compact operator on a separable Hilbert space $\mathcal{H}$. If the positive eigenvalues of A are listed in decreasing order as $\lambda_1 \geq \lambda_2 \geq \ldots$, then*

$$\lambda_k = \max_{W \in \Lambda_k} \left\{ \min_{u \in W \setminus \{0\}} \frac{\langle u, Au \rangle}{\|u\|^2} \right\},$$

where Λ_k denotes the set of subspaces of $\mathcal{H}$ of dimension k.

Proof Let $\{\phi_k\}$ denote the orthonormal basis vectors corresponding to the eigenvalues $\{\lambda_k\}$. For $u \in \operatorname{span}\{\phi_1, \ldots, \phi_k\}$, the basis expansion implies that

$$\begin{aligned} \langle u, Au \rangle &= \sum_{j=1}^{k} \lambda_j |\langle u, \phi_j \rangle|^2 \\ &\geq \lambda_k \|u\|^2 \end{aligned}$$

(because the eigenvalues are decreasing). Therefore

$$\min_{u \in \operatorname{span}\{\phi_1, \ldots, \phi_k\} \setminus \{0\}} \frac{\langle u, Au \rangle}{\|u\|^2} \geq \lambda_k.$$

On the other hand, for an arbitrary $W \in \Lambda_k$, $W \cap \operatorname{span}\{\phi_1, \ldots, \phi_{k-1}\}^\perp$ has dimension at least 1. Therefore we can choose a nonzero vector $w \in W \cap \operatorname{span}\{\phi_1, \ldots, \phi_{k-1}\}^\perp$. The fact that $w \in \operatorname{span}\{\phi_1, \ldots, \phi_{k-1}\}^\perp$ implies that w is a combination of eigenvectors with eigenvalues less than or equal to λ_k. Thus

$$\langle w, Aw \rangle \leq \lambda_k \|w\|^2,$$

implying that

$$\min_{u \in W \setminus \{0\}} \frac{\langle u, Au \rangle}{\|u\|^2} \leq \lambda_k.$$

□

4.4.2 Hilbert–Schmidt Operators

Suppose that A is a compact self-adjoint operator on $L^2(\Omega)$, where Ω is an open subset of $\mathbb{R}^n$. In Section 3.5, we introduced the notion of a Hilbert–Schmidt operator,

which can be represented with an L^2 integral kernel. In terms of the eigenvalues and eigenvectors given by Theorem 4.21, this kernel has the form,

$$K(x, y) = \sum_{k=1}^{\infty} \lambda_k \phi_k(x)\overline{\phi_k(y)}, \tag{4.33}$$

which converges in $L^2(\Omega \times \Omega)$. For certain applications, it is useful to be able to strengthen the convergence of (4.33) to uniform convergence by placing additional assumptions on A and K.

Theorem 4.23 (Mercer's Theorem). *Suppose $\Omega \subset \mathbb{R}^n$ is a bounded domain and A is a positive Hilbert–Schmidt operator on $L^2(\Omega)$. If the integral kernel $K(\cdot, \cdot)$ is continuous on $\Omega \times \Omega$, then the eigenfunction ϕ_k is continuous on Ω if $\lambda_k > 0$, and the expansion (4.33) converges uniformly on compact sets.*

Proof Note that it suffices to assume that $\lambda_k > 0$ for all k, since terms with $\lambda_k = 0$ do not affect the sum (4.33). By the definition of K and the eigenvalue equation,

$$\phi_k(x) = \frac{1}{\lambda_k} \int_\Omega K(x, y)\phi_k(y)\, dy. \tag{4.34}$$

Since Ω is bounded, the eigenfunction ϕ_k is also in $L^1(\Omega)$, by Fubini's theorem. Therefore the continuity of ϕ_k follows from (4.34) by the dominated convergence theorem.

We claim that the positivity of A implies that $K(x, x) \geq 0$ for all $x \in \Omega$. To see this, note that if $K(x_0, x_0) < 0$ for some $x_0 \in \Omega$, then by continuity $K(\cdot, \cdot) < 0$ on $U \times U$ for some neighborhood U of x_0. This would imply

$$\langle \chi_U, A\chi_U \rangle < 0,$$

contradicting the positivity of A.

For $N \in \mathbb{N}$, let us define the partial sum

$$K_N(x, y) := \sum_{k=1}^{N} \lambda_k \phi_k(x)\overline{\phi_k(y)},$$

with the remainder $R_N(x, y) = K(x, y) - K_N(x, y)$. Both K_N and R_N are continuous. From the L^2 expansion,

$$R_N(x, y) := \sum_{k=N+1}^{\infty} \lambda_k \phi_k(x)\overline{\phi_k(y)},$$

we can see that $R_N(x, y)$ is the kernel of a positive operator. Therefore $R_N(x, x) \geq 0$, by the same reasoning applied to K.

We conclude that $K_N(x,x) \le K(x,x)$ for all N, which then implies

$$\sum_{k=1}^{\infty} \lambda_k |\phi_k(x)|^2 \le K(x,x). \tag{4.35}$$

Dini's theorem from real analysis (see, e.g., Rudin [77, Thm. 7.13]) says that if a monotonic sequence of continuous functions on a compact set converges pointwise to a continuous function, then the convergence is uniform. Therefore, the bound (4.35) implies that $\sum \lambda_k |\phi_k(x)|^2$ converges uniformly on compact sets.

To extend this result off the diagonal, note that the Cauchy–Schwarz inequality gives

$$\left| \sum_{k=m_1}^{m_2} \lambda_k \phi_k(x) \overline{\phi_k(y)} \right|^2 \le \sum_{k=m_1}^{m_2} \lambda_k |\phi_k(x)|^2 \sum_{k=m_1}^{m_2} \lambda_k |\phi_k(y)|^2.$$

Hence, the series

$$\sum_{k=1}^{\infty} \lambda_k \phi_k(x) \overline{\phi_k(y)}$$

converges uniformly on compact sets. The limit is continuous, and therefore equal to K because (4.33) holds in the L^2 sense. □

4.4.3 Traces

Suppose that A is a compact self-adjoint operator on a separable Hilbert space $\mathcal{H}$. By Theorem 4.21 there exists an orthonormal basis $\{\phi_k\}$ consisting of eigenvectors such that the corresponding eigenvalues $\{\lambda_k\}$ are discrete with $\lambda_k \to 0$. The operator A is said to be *trace-class* if

$$\sum_k |\lambda_k| < \infty,$$

in which case we define

$$\operatorname{tr} A := \sum_{k=1}^{\infty} \lambda_k. \tag{4.36}$$

A compact operator T is Hilbert–Schmidt if and only if T^*T is trace-class, and the abstract Hilbert–Schmidt norm introduced in Section 3.5 can be computed as

$$\|T\|_{\mathrm{HS}}^2 = \operatorname{tr} T^*T.$$

In many spectral theory applications, the compact operator is given by an integral kernel, and it is particularly useful to be able to express the trace as an integral over this kernel.

Theorem 4.24. *Suppose $\Omega \subset \mathbb{R}^n$ is a bounded domain. If A is a positive Hilbert–Schmidt operator on $L^2(\Omega)$, with continuous kernel $K(\cdot, \cdot)$, then*

$$\operatorname{tr} A = \int_\Omega K(x,x)\, dx,$$

where A is trace-class if and only if the integral is finite.

Proof Let $\{\lambda_k\}$ and $\{\phi_k\}$ denote the eigenvalues and eigenfunctions of A. Since $\lambda_k \geq 0$, the monotone convergence theorem implies that

$$\sum_{k=1}^\infty \lambda_k = \int_\Omega \sum_{k=1}^\infty \lambda_k |\phi_k(x)|^2\, dx.$$

If K is continuous, it then follows from Mercer's theorem (Theorem 4.23) that

$$\sum_{k=1}^\infty \lambda_k = \int_\Omega K(x,x)\, dx.$$

□

4.5 Exercises

4.1. Prove the second resolvent identity: If S and T are operators with $\mathcal{D}(S) = \mathcal{D}(T)$, then for $z \in \rho(S) \cap \rho(T)$,

$$(S-z)^{-1} - (T-z)^{-1} = (S-z)^{-1}(S-T)(T-z)^{-1}.$$

4.2. Using the definition (4.17) for the derivative of an operator-valued function, prove that, for $z \in \rho(T)$,

$$\frac{d^n}{dz^n}(T-z)^{-1} = n!(T-z)^{-n}.$$

4.3. Let A be a closed operator on $\mathcal{H}$ and suppose $z \in \rho(A)$. If B is a bounded operator, prove that $z \in \rho(A+B)$ for $\|B\|$ sufficiently small.

4.4. Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal basis for $\mathcal{H}$, and define the shift operator T by

$$Te_k = e_{k+1}$$

for $k \in \mathbb{N}$.

(a) Show that T has no eigenvalues.
(b) Determine the eigenvalues of T^*.
(c) Compute $\sigma(T)$.

4.5. Let U be a unitary operator on a Hilbert space $\mathcal{H}$. Prove that the spectrum of U is contained in the unit circle.

4.6. Prove that the spectrum of a positive self-adjoint operator A satisfies $\sigma(A) \subset [0, \infty)$.

4.7. Suppose that $q(x)$ is a polynomial.

(a) For $T \in \mathcal{L}(\mathcal{H})$, prove that

$$\sigma(q(T)) = q(\sigma(T)).$$

(b) For a self-adjoint operator $A \in \mathcal{L}(\mathcal{H})$, prove that

$$\|q(A)\| = \sup_{\lambda \in \sigma(A)} |q(\lambda)|.$$

4.8. (Continuous Functional Calculus) Assume that $A \in \mathcal{L}(\mathcal{H})$ is self-adjoint and $f \in C(\sigma(A))$.

(a) By the Stone–Weierstrass theorem, there exists a sequence of polynomials $\{p_n(x)\}$ such that $p_n \to f$ uniformly on $\sigma(A)$. Use Exercise 4.7 to show that the limit

$$f(A) := \lim_{n \to \infty} p_n(A)$$

exists (with respect to the operator norm) and is independent of the choice of polynomials.
(b) Prove that

$$\|f(A)\| = \|f\|_\infty.$$

(c) Prove that

$$\sigma(f(A)) = f(\sigma(A)).$$

4.9. Suppose that A is a symmetric operator on $\mathcal{H}$ and that $\mathcal{H}$ has an orthonormal basis $\{\phi_k\}_{k=1}^{\infty}$ with $\phi_k \in \mathcal{D}(A)$ and

$$A\phi_k = \lambda_k \phi_k$$

for $\lambda_k \in \mathbb{R}$.

(a) Prove that A is essentially self-adjoint.
(b) Show that $\sigma(\overline{A})$ is the closure of $\{\lambda_k\}$.

4.10. On $L^2(0, 1)$, consider the *Volterra operator*

$$Tf(x) = \int_0^x f(t)\, dt.$$

(a) Show that T is a Hilbert–Schmidt integral operator and therefore compact.
(b) Show that $\sigma(T) = \{0\}$.
(c) Compute T^* explicitly.
(d) Find the eigenvalues of T^*T and use this to determine $\|T\|$.

4.11. Suppose that the operator T has *compact resolvent*, which means that $(T - z_0)^{-1}$ is compact for some $z_0 \in \rho(T)$.

(a) Prove that $(T - z)^{-1}$ is compact for all $z \in \rho(T)$.
(b) Prove that $\sigma(T)$ is discrete as a subset of $\mathbb{C}$.

4.12. For a closed operator T and $\varepsilon > 0$, the *ε-pseudospectrum* of T is defined as

$$\Sigma_\varepsilon(T) := \sigma(T) \cup \left\{ z \in \rho(T) : \left\| (T - z)^{-1} \right\| > 1/\varepsilon \right\}.$$

Prove that

$$\Sigma_\varepsilon(T) = \bigcup_{B \in \mathcal{L}(\mathcal{H}):\, \|B\| < \varepsilon} \sigma(T + B).$$

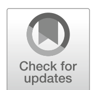
Notes

The basic definitions of resolvent and spectrum extend to the case of bounded operators on Banach spaces and play an important role in the theory of C^*-algebras. Many properties that we have proven for operators on Hilbert spaces, such as analyticity of the resolvent and the spectral radius formula, extend to the

Banach space context. Our discussion of these topics draws inspiration from Reed and Simon [69, §VI.3], but the details are simplified for Hilbert space operators. Additional background on the spectrum of operators on Banach spaces can be found in many texts; see, e.g., Arveson [3, Chapter 1], MacCluer [60, Chapter 5], or Simon [83, Chapter 2].

Our approach to the spectral theory of compact operators, based on the analytic Fredholm theorem, follows Reed and Simon [69, §VI.5]. Another standard method is to first show that a compact self-adjoint operator A has either $\pm\|A\|$ as an eigenvalue, and then argue inductively. See, e.g., Stein and Shakarchi [87, §4.6] or MacCluer [60, §4.3].

Chapter 5
The Spectral Theorem



The first spectral theorem for matrices was proven by Augustin-Louis Cauchy, who established that a real symmetric is diagonalizable in 1826. Charles Hermite extended this result in 1855, proving that a complex self-adjoint $n \times n$ matrix has n real eigenvalues, and there exists an orthonormal basis for $\mathbb{C}^n$ consisting of eigenvectors. The Hilbert–Schmidt theorem (Theorem 4.21) shows that compact self-adjoint operators are diagonalizable in the same sense.

In the theory of operators, the role of diagonal matrices is played by multiplication operators of the type discussed in Example 3.2. The operator version of the spectral theorem says that a self-adjoint operator is unitarily equivalent to a multiplication operator. This was established independently by Marshall Stone and John von Neumann in the early 1930s. Their development of the spectral theorem was motivated by quantum mechanics, where self-adjoint operators play a central role as the representations of physical observables such as position, momentum, and energy.

We have already seen a special case of the operator spectral theorem in Example 3.23, where we noted that the Fourier transform on $\mathbb{R}^n$ conjugates the Laplacian to multiplication by $|\xi|^2$. To illustrate the difference between the multiplication operator form of the spectral theorem and the matrix version, let us consider the case of a diagonal operator.

Example 5.1. Suppose that A is a self-adjoint operator on $\mathcal{H}$, with an orthonormal basis of eigenvectors $\{\phi_n\}$ such that $A\phi_n = \lambda_n \phi_n$ for $\lambda_n \in \mathbb{R}$. Assume furthermore that the eigenvalues λ_n are distinct. To express A as a multiplication operator, we define the measure

$$\mu := \sum_n \delta_{\lambda_n}, \tag{5.1}$$

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1_5

where δ_x denotes the point measure at $x \in \mathbb{R}$. The unitary map $Q : L^2(\mathbb{R}, d\mu) \to \mathcal{H}$ defined by

$$Qf := \sum_n f(\lambda_n)\phi_n \tag{5.2}$$

conjugates A to a simple multiplication operator,

$$Q^{-1}AQ = M_x. \tag{5.3}$$

Note that if we applied the same construction to an eigenvalue λ with multiplicity m, the corresponding term in (5.1) would be $m\delta_\lambda$. This changes the normalization of the measure, requiring some adjustment to (5.2) to make Q unitary, but it does not affect the multiplicity of λ as an eigenvalue of M_x. To accommodate higher multiplicities in this construction, we would need to take multiple copies of $\mathbb{R}$. $\lozenge$

The spectral theorem can be extended from the self-adjoint case to the class of *normal operators*, i.e., bounded operators which commute with their adjoints. In fact, the extension to normal operators could be derived as a special case of a joint spectral theorem for commuting families of bounded operators. We will limit our attention to the unbounded self-adjoint case in this chapter, because that is the relevant context for all of the applications discussed later in the book. Our approach to the proof involves exploiting the connection between self-adjoint and unitary operators, a trick due to von Neumann.

5.1 Unitary Operators

In operator theory, the term "functional calculus" refers to the ability to apply a function to an operator. One possible construction of the continuous functional calculus for bounded self-adjoint operators was developed in Exercises 4.7 and 4.8, based on polynomial approximations.

In this section we will give a derivation of the functional calculus for unitary operators based on Fourier series, and use it to prove a spectral theorem. Recall from (3.6) that a map $U \in \mathcal{L}(\mathcal{H})$ is unitary if and only if

$$U^*U = UU^* = I.$$

This formula suggests that the spectrum of a unitary operator should be a subset of the unit circle $\mathbb{S} := \{|z| = 1\} \subset \mathbb{C}$, which is indeed the case (see Exercise 4.5).

5.1.1 Continuous Functional Calculus

Our first goal is to define $f(U)$ for U unitary and f a continuous function $\mathbb{S} \to \mathbb{C}$. If f has a convergent expansion with respect to the Fourier basis $\{e^{ik\theta}\}$, then we can define $f(U)$ by replacing $e^{ik\theta}$ in the series with U^k. This will provide the first step in our construction.

For a function $f \in L^1(\mathbb{S})$, define the discrete Fourier coefficients,

$$\hat{f}(k) := \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) e^{-ik\theta}\, d\theta \tag{5.4}$$

for $k \in \mathbb{Z}$. If $f \in C^\infty(\mathbb{S})$, then repeated integration by parts gives the estimate

$$\hat{f}(k) = O\big((1 + |k|)^{-n}\big) \tag{5.5}$$

for all $n \in \mathbb{N}$. In this case, we define $f(U)$ for $U \in \mathcal{L}(\mathcal{H})$ unitary by the convergent series

$$f(U) := \sum_{k \in \mathbb{Z}} \hat{f}(k) U^k. \tag{5.6}$$

To extend this definition to continuous functions, the crucial fact to establish is the positivity of $f(U)$ when $f \geq 0$. This will allow us to control the convergence $f_n(U) \to f(U)$ in the operator topology, assuming that the sequence $\{f_n\} \subset C^\infty(\mathbb{S})$ converges uniformly to f. Recall from Section 3.4 that an operator A is *positive* ($A \geq 0$) if

$$\langle v, Av \rangle \geq 0, \tag{5.7}$$

for all $v \in \mathcal{D}(A)$.

Theorem 5.2 (Continuous Functional Calculus for Unitary Operators). *Suppose $U \in \mathcal{L}(\mathcal{H})$ is unitary. The map $f \mapsto f(U)$ defined by (5.6) extends uniquely to a continuous map $C(\mathbb{S}) \to \mathcal{L}(\mathcal{H})$ with the following properties:*

(a) $f(U)^* = \overline{f}(U)$.
(b) $f(U)g(U) = (fg)(U)$.
(c) *If $f \geq 0$, then $f(U) \geq 0$.*
(d) $\|f(U)\| \leq \sup |f|$.

Proof We first establish the properties in the smooth case, $f \in C^\infty(\mathbb{S})$. Since $U^* = U^{-1}$, taking the adjoint of (5.6) gives

$$f(U)^* = \sum_{k \in \mathbb{Z}} \overline{\hat{f}(k)} U^{-k}. \tag{5.8}$$

By the definition of the coefficients (5.4),

$$\overline{\hat{f}(k)} = \widehat{\overline{f}}(-k),$$

so (5.8) proves (a) for $f \in C^\infty(\mathbb{S})$. For $f, g \in C^\infty(\mathbb{S})$, property (b) follows immediately from the convolution formula

$$\widehat{fg}(k) = \sum_{l \in \mathbb{Z}} \hat{f}(l)\hat{g}(k-l).$$

To prove (c) in the case $f \in C^\infty(\mathbb{S})$, $f \geq 0$, consider the function

$$h_\varepsilon(z) := \sqrt{f(z) + \varepsilon}$$

for $\varepsilon > 0$. This is smooth, so $h_\varepsilon(U)$ is defined by (5.6) and self-adjoint by (a). Furthermore,

$$h_\varepsilon(U)^2 = f(U) + \varepsilon$$

by (b). For $v \in \mathcal{H}$, we thus have

$$\begin{aligned} \langle v, f(U)v \rangle &= \left\langle v, \left(h_\varepsilon(U)^2 - \varepsilon\right)v \right\rangle \\ &= \|h_\varepsilon(U)v\|^2 - \varepsilon\|v\|^2 \\ &\geq -\varepsilon\|v\|^2. \end{aligned}$$

Taking $\varepsilon \to 0$ yields $\langle v, f(U)v \rangle \geq 0$.

Property (d) is a simple consequence of (c). If $f \in C^\infty(\mathbb{S})$ and $M := \sup|f|$, then $M^2 - |f|^2 \geq 0$. Hence, by (c),

$$\left\langle v, (M^2 - |f|^2(U))v \right\rangle \geq 0.$$

Since $\langle v, |f|^2(U)v \rangle = \|f(U)v\|^2$ by (a) and (b), we deduce that

$$\|f(U)v\| \leq M\|v\|.$$

This completes the proof of (a)–(d) in the smooth case. The extension to $C(\mathbb{S})$ is now straightforward. By the Weierstrass approximation theorem, $C^\infty(\mathbb{S})$ is dense in $C(\mathbb{S})$ with respect to the sup norm. Given $f \in C(\mathbb{S})$, choose a sequence $\{f_n\} \subset C^\infty(\mathbb{S})$ such that $f_n \to f$ uniformly on $\mathbb{S}$. The corresponding operator sequence $\{f_n(U)\}$ is Cauchy in $\mathcal{L}(\mathcal{H})$ by (d), and therefore we can define $f(U) := \lim f_n(U)$. It is easy to check that $f(U)$ is independent of the choice of $\{f_n\}$, and that the properties (a)–(d) are preserved in the limit. □

5.1.2 Spectral Measures

The connection between the continuous functional calculus of Theorem 5.2 and the multiplication operator form of the spectral theorem is provided by a fundamental result from measure theory. We will use the functional calculus itself to construct the measure, by considering functionals $C(\mathbb{S}) \to \mathbb{C}$ of the form

$$f \mapsto \langle v, f(U)v \rangle, \tag{5.9}$$

for $v \in \mathcal{H}$. By property (c) of Theorem 5.2, this functional is positive in the sense that $f \geq 0$ implies

$$\langle v, f(U)v \rangle \geq 0.$$

A positive functional on $C(\mathbb{S})$ yields a Borel measure on $\mathbb{S}$, by the following:

Theorem 5.3 (Riesz Representation Theorem). *Let X be a compact metric space. Given a positive linear functional $\beta : C(X) \to \mathbb{C}$, there exists a unique Borel measure μ on X such that*

$$\beta(f) = \int_X f \, d\mu$$

for $f \in C(X)$.

We will only sketch the proof of Riesz representation here; the details are reviewed in Appendix A.1.6. Given a positive functional β, we define the measure of an open subset $\Omega \subset X$ by

$$\mu_0(\Omega) := \sup \left\{ \beta(f) : \ f \in C(X), \ 0 \leq f \leq 1, \operatorname{supp} f \subset \Omega \right\}$$

The measure of a Borel set $A \subset X$ is then given by

$$\mu(A) := \inf \left\{ \mu_0(\Omega) : \ A \subset \Omega \text{ with } \Omega \text{ open} \right\}.$$

Let $U \in \mathcal{L}(\mathcal{H})$ be unitary. For $v \in \mathcal{H}$, Theorem 5.2(c) implies that linear functional (5.9) is positive on $C(\mathbb{S})$. By Theorem 5.3, there is an associated Borel measure μ_v on $\mathbb{S}$ such that

$$\langle v, f(U)v \rangle = \int_{\mathbb{S}} f \, d\mu_v. \tag{5.10}$$

This is called the *spectral measure* associated with v.

Lemma 5.4. *Given U and v as above, the map $W_v : f \mapsto f(U)v$ has a unique continuous extension to an isometry*

$$W_v : L^2(\mathbb{S}, d\mu_v) \to \mathcal{H},$$

such that

$$UW_v[f(z)] = W_v[zf(z)]. \tag{5.11}$$

Proof Let $f, g \in C(\mathbb{S})$. By the definition of W_v and Theorem 5.2,

$$\begin{aligned}\langle W_v f, W_v g\rangle &= \big\langle f(U)v, g(U)v\big\rangle \\ &= \big\langle v, f(U)^* g(U)v\big\rangle \\ &= \big\langle v, \overline{f} g(U)v\big\rangle.\end{aligned}$$

Thus, by the definition (5.10),

$$\begin{aligned}\langle W_v f, W_v g\rangle &= \int_{\mathbb{S}} \overline{f} g \, d\mu_v \\ &= \langle f, g\rangle_{L^2(\mathbb{S}, d\mu_v)}.\end{aligned}$$

It follows easily from the Riesz construction that $C(\mathbb{S})$ is dense in $L^2(\mathbb{S}, d\mu_v)$; see Lemma A.11. Since W_v preserves the inner product for functions in $C(\mathbb{S})$, it follows that W_v has a unique continuous extension to an isometry $L^2(\mathbb{S}, d\mu_v) \to \mathcal{H}$.

The identity (5.11) follows from the fact that $(zf)(U) = Uf(U)$ by Theorem 5.2(b). □

5.1.3 Spectral Theorem for Unitary Operators

The construction of spectral measures leads us directly to the proof of the spectral theorem in the unitary case. To set up the statement, let us recall some basic features of multiplication operators from Example 3.5 and Section 4.1.2. Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space. For a multiplication operator M_f acting on $L^2(X, d\mu)$ with

$$\mathcal{D}(M_f) = \big\{u \in L^2(X, d\mu) : fu \in L^2(X, d\mu)\big\},$$

the adjoint is given by $M_f^* = M_{\overline{f}}$ on the same domain $\mathcal{D}(M_f^*) = \mathcal{D}(M_f)$. Thus, M_f is unitary if and only if $|f| = 1$ almost everywhere with respect to μ.

For $v \in \mathcal{H}$ it is clear from (5.6) that the range of the map W_v from Lemma 5.4 is the space

$$H_v := \overline{\operatorname{span}\big\{U^k v, k \in \mathbb{Z}\big\}}.$$

It is possible that $H_v = \mathcal{H}$, in which case the unitary form of the spectral theorem is already proven by Lemma 5.4. If H_v does not cover all of $\mathcal{H}$, we can iterate the construction to obtain the following:

Theorem 5.5 (Spectral Theorem for Unitary Operators). *Suppose $\mathcal{H}$ is a separable Hilbert space and $U \in \mathcal{L}(\mathcal{H})$ is unitary. Then there exists a measure space (Y, ν), defined as*

$$(Y, \nu) = \cup_k (\mathbb{S}, \nu_k),$$

where $\{\nu_k\}$ is a sequence of finite measures, and a unitary map

$$W : L^2(Y, d\nu) \to \mathcal{H}$$

such that

$$W^{-1} U W = M_\eta, \tag{5.12}$$

where $\eta(z) := z$ on each copy of $\mathbb{S}$.

Proof Let $\{w_j\}$ be a countable dense subset of $\mathcal{H}$. Applying Lemma 5.4 to the vector w_1 gives a measure ν_1 and an isomorphism

$$W_1 : L^2(\mathbb{S}, d\nu_1) \to H_1.$$

If $H_1 = \mathcal{H}$, then this completes the proof. Otherwise, note that U preserves H_1 by definition, and by unitarity U also preserves $H_1^\perp$.

Next pick the first j such that $w_j \notin H_1$, and let v_2 be the orthogonal projection of this w_j into $H_1^\perp$. Applying Lemma 5.4 to v_2 yields a measure μ_2 and an isometry W_2 with range H_2. Then, either $H_1 \oplus H_2 = \mathcal{H}$, or else we take the first $w_j \notin H_1 \oplus H_2$ and use this to define $v_3 \in (H_1 \oplus H_2)^\perp$ by orthogonal projection.

Continuing this process yields a sequence $\{v_k\}$ (possibly finite) such that $w_j \in \oplus_k H_k$ for all j. It follows that

$$\mathcal{H} = \oplus_k H_k.$$

(See Section 2.4 for the definition of a countable direct sum.) For the measure space $(Y, \nu) := \cup_k (\mathbb{S}, \nu_k)$, there is a corresponding decomposition

$$L^2(Y, d\nu) = \oplus_k L^2(\mathbb{S}, d\nu_k).$$

We thus have a unitary map $W : L^2(Y, d\nu) \to \mathcal{H}$ given by $\oplus_k W_k$. The relation (5.12) follows from (5.11). □

5.2 The Main Theorem

We are now prepared to state and prove the spectral theorem for self-adjoint operators by means of the unitary version.

Theorem 5.6 (Spectral Theorem—Multiplication Operator Form). *Suppose A is a self-adjoint operator on a separable Hilbert space $\mathcal{H}$. There exists a countable collection of finite Borel measures $\{\mu_k\}$ on $\mathbb{R}$ and a unitary map $Q : L^2(X, \mu) \to \mathcal{H}$, with $(X, \mu) = \cup_k(\mathbb{R}, \mu_k)$, such that*

$$Q^{-1}AQ = M_\alpha$$

and

$$\mathcal{D}(A) = Q\mathcal{D}(M_\alpha),$$

where $\alpha : X \to \mathbb{R}$ is given by $\alpha(x) := x$ on each copy of $\mathbb{R}$.

Note that the unitary equivalence of A with M_α implies, by Theorem 4.5, that

$$\sigma(A) = \text{ess-range}(\alpha),$$

where the essential range is defined with respect to μ. This means that the measures μ_k have support within $\sigma(A)$.

The remainder of this section is devoted to the proof of Theorem 5.6. For bounded self-adjoint operators, it is possible to give a proof analogous to that of Theorem 5.5, by developing the continuous functional calculus directly as in Exercise 4.8. The advantage of the unitary operator approach is that it allows us to treat bounded and unbounded self-adjoint operators on an equal footing.

The association between self-adjoint and unitary operators is inspired by the *Cayley transformation*

$$\gamma(x) := \frac{x - i}{x + i}, \tag{5.13}$$

which maps the real line to the unit circle $\mathbb{S} \subset \mathbb{C}$.

Lemma 5.7 (Cayley Transform). *If A is self-adjoint, then*

$$U := I - 2i(A + i)^{-1}$$

is unitary.

Proof By Theorem 4.4 and the fact that A is self-adjoint,

$$\left[(A + i)^{-1}\right]^* = (A - i)^{-1}.$$

Therefore,

$$U^* = I + 2i(A - i)^{-1},$$

and U commutes with U^* because $(A - i)^{-1}$ and $(A + i)^{-1}$ commute.

To complete the proof, note that

$$U^*U = I + 2i(A-i)^{-1} - 2i(A+i)^{-1} + 4(A-i)^{-1}(A+i)^{-1}.$$

By the first resolvent formula (Corollary 4.12),

$$2i(A-i)^{-1} - 2i(A+i)^{-1} + 4(A-i)^{-1}(A+i)^{-1} = 0.$$

Therefore $U^*U = I$, and so U is unitary. □

The Cayley transform allows many results to be translated between the unitary and self-adjoint cases. In particular, we will now use it to show that the spectral theorem for self-adjoint operators follows from Theorem 5.5.

Proof of Theorem 5.6 Given A self-adjoint, let U by the corresponding unitary operator defined by the Cayley transform. From Theorem 5.5 we obtain the decomposition $\mathcal{H} = \oplus H_k$, with a corresponding sequence of finite measures ν_k on $\mathbb{S}$ and unitary maps $W_k : L^2(\mathbb{S}, \nu_k) \to H_k$, such that

$$W_k^{-1} U W_k = M_z, \tag{5.14}$$

where z is the complex coordinate on $\mathbb{S}$.

The operator $I - U = 2i(A+i)^{-1}$ is injective on $\mathcal{H}$. Hence M_{1-z} is injective on each component of $L^2(\mathbb{S}, \nu_k)$, by (5.14). It follows that $\nu_k\{1\} = 0$ for each k. We can therefore use the inverse Cayley map,

$$\eta(z) := i\frac{1+z}{1-z},$$

to define the push-forward of ν_k to a finite Borel measure on $\mathbb{R}$,

$$\mu_k := \eta_* \nu_k.$$

There is a corresponding unitary map $\Psi_k : L^2(\mathbb{R}, d\mu_k) \to L^2(\mathbb{S}, d\nu_k)$ given by the pullback $\phi \mapsto \phi \circ \eta$.

Let $(X, \mu) = \cup_k (\mathbb{R}, \mu_k)$, and define the unitary map

$$Q : L^2(X, d\mu) \to \mathcal{H}$$

by $Q = \oplus_k Q_k$ where

$$Q_k := W_k \Psi_k : L^2(\mathbb{R}, d\mu_k) \to H_k.$$

On X we define the coordinate function $\alpha : X \to \mathbb{R}$ which is given by $\alpha_k(x) = x$ on each copy of $\mathbb{R}$ indexed by k.

To prove that $Q(\mathcal{D}(M_\alpha)) \subset \mathcal{D}(A)$, it suffices to consider each copy of $\mathbb{R}$ separately. Suppose that $f \in \mathcal{D}(M_x) \subset L^2(\mathbb{R}, d\mu_k)$. This implies in particular that $g(x) := (x+i)f(x)$ lies in $L^2(\mathbb{R}, d\mu_k)$. Applying the Cayley pullback gives

$$\begin{aligned}[][\Psi_k g](z) &= (\eta(z)+i) f \circ \eta(z) \\ &= \frac{2i}{1-z} f \circ \eta(z).\end{aligned}$$

Therefore, by (5.14),

$$(I-U)Q_k g = 2i Q_k f.$$

This means that

$$Q_k f = (A+i)^{-1} Q_k g,$$

which shows that $Q_k f \in \mathcal{D}(A)$. The same argument applies for any k, hence

$$Q(\mathcal{D}(M_\alpha)) \subset \mathcal{D}(A).$$

Now suppose that $v \in H_k \cap \mathcal{D}(A)$, and set $w := (A+i)v$. From

$$v = (A+i)^{-1} w = \frac{1}{2i}(I-U)w$$

and (5.14) we derive that

$$W_k^{-1} v = \frac{1}{2i} M_{1-z} W_k^{-1} w.$$

Therefore

$$\begin{aligned} M_x Q_k^{-1} v &= \frac{1}{2i} \Psi_k^{-1} M_{\eta(z)} M_{1-z} W_k^{-1} w \\ &= \frac{1}{2} \Psi_k^{-1} M_{1+z} W_k^{-1} w \\ &= \frac{1}{2} Q_k^{-1} (1+U) w. \end{aligned} \tag{5.15}$$

It follows that $M_x Q_k^{-1} v \in L^2(\mathbb{R}, d\mu_k)$, hence $v \in Q_k(\mathcal{D}(M_x))$. Applying this argument to each component gives

$$\mathcal{D}(A) \subset Q(\mathcal{D}(M_\alpha)),$$

completing the proof that $Q(\mathcal{D}(M_\alpha)) = \mathcal{D}(A)$.

Finally, from (5.15) we deduce that on $\mathcal{D}(A)$,

$$\begin{aligned} Q M_\alpha Q^{-1} &= \frac{1}{2}(I+U)(A+i) \\ &= A. \end{aligned}$$

□

Since multiplicity problem makes the notation used in the proof of the spectral theorem a bit confusing, let us illustrate the construction in a simple case with trivial multiplicity.

Example 5.8. Consider the operator $A = M_x$ on $L^2(\mathbb{R})$. The Cayley transform of A is $U = M_{\gamma(x)}$, where γ was defined in (5.13). This operator admits a cyclic unit vector $v \in L^2(\mathbb{R})$ given by

$$v(x) = \frac{1}{\sqrt{\pi}} \frac{1}{x+i}.$$

Given $f \in C(\mathbb{S})$, we have

$$\begin{aligned} \langle v, f(U)v \rangle &= \int_{\mathbb{R}} f(\gamma(x)) \, \frac{\pi}{x^2+1} \, dx \\ &= \int_0^{2\pi} f(e^{i\theta}) \, \frac{d\theta}{2\pi}. \end{aligned}$$

In this case the measure produced by Theorem 5.5 is the standard circle measure $d\nu = d\theta/2\pi$. The unitary map $W : L^2(\mathbb{S}, d\nu) \to \mathcal{H}$ is defined by $f \mapsto f(U)v$, and so is given by

$$[Wf](x) = f(\gamma(x))v(x).$$

The measure $\mu = \eta_* \nu$, where η is the inverse Cayley map, is given by

$$d\mu = \frac{1}{\pi} \frac{dx}{x^2+1}.$$

The corresponding unitary map $\Psi : L^2(\mathbb{R}, d\mu) \to L^2(\mathbb{S}, d\nu)$ is the pullback $\Psi : \phi \mapsto \phi \circ \eta$. Thus $Q := W\Psi$ is given simply by

$$Q = M_v,$$

which is unitary as a map $L^2(\mathbb{R}, d\mu) \to L^2(\mathbb{R})$. As one might expect, since we started from a multiplication operator, the conjugation of A by Q is trivial, $Q^{-1} M_x Q = M_x$. ♢

5.3 Functional Calculus

We have already made use of a continuous form of the functional calculus for unitary operators in Section 5.1. One of the principal applications of Theorem 5.6 is a functional calculus for self-adjoint operators which includes the broader class of Borel functions. We can use it, for example, to define projections that isolate different parts of the spectrum. There are many applications for the functional calculus in PDE theory, in creating solution operators for equations. For example, the heat operator $e^{t\Delta}$ maps initial data to a solution of the heat equation on $\mathbb{R}^n$.

The functional calculus developed in this section and the multiplication operator form of the spectral theorem stated in Theorem 5.6 are essentially equivalent. At least in the bounded case, one can quickly derive the multiplication operator form from the functional calculus by means of the Riesz representation theorem on $\mathbb{R}$. The corresponding argument for unbounded operators is less direct, however, which is why we have taken the multiplication operator form as the primary version of the spectral theorem.

The *Borel functions* on $\mathbb{R}$ are the complex-valued functions for which the preimage of a Borel set is a Borel set. (See Appendix A.1 for background measure theory definitions.) By the definition of measurability, if $g : X \to \mathbb{R}$ is measurable and $f : \mathbb{R} \to \mathbb{C}$ is Borel, then the composition $f \circ g$ is measurable.

Let $\mathcal{B}_b(\mathbb{R})$ denote the space of bounded Borel functions $\mathbb{R} \to \mathbb{C}$. Given a self-adjoint operator A on $\mathcal{H}$ and $f \in \mathcal{B}_b(\mathbb{R})$, we define the bounded operator

$$f(A) := QM_{f\circ\alpha}Q^{-1}, \tag{5.16}$$

where $Q : L^2(X, d\mu) \to \mathcal{H}$ and $\alpha : X \to \mathbb{R}$ are defined as in Theorem 5.6.

Theorem 5.9 (Functional Calculus). *For a self-adjoint operator A, the map $\mathcal{B}_b(\mathbb{R}) \to \mathcal{L}(\mathcal{H})$ defined by $f \mapsto f(A)$ has the following properties:*

(a) *The map is a $*$-homomorphism, meaning*

$$fg(A) = f(A)g(A), \qquad f(A)^* = \overline{f}(A).$$

(b) *For $f \in \mathcal{B}_b(\mathbb{R})$,*

$$\|f(A)\| \le \sup_{\lambda\in\sigma(A)} |f(\lambda)|, \tag{5.17}$$

with equality if f is continuous.

(c) *If $f_n \to f$ pointwise and $\sup|f_n| \le M$ for all n, then $f_n(A) \to f(A)$ in the strong operator sense, i.e., $f_n(A)v \to f(A)v$ for all $v \in \mathcal{H}$.*

Moreover, (5.16) gives the unique map $\mathcal{B}_b(\mathbb{R}) \to \mathcal{L}(\mathcal{H})$ satisfying these conditions.

Proof Property (a) follows directly from the corresponding results for multiplication operators, by means of the equivalence (5.16). The adjoint property was covered in Exercise 3.12.

In Theorem 4.5 we proved that the spectrum of the multiplication operator M_α is the essential range of α. Therefore, (5.16) implies that

$$\sigma(A) = \text{ess-range}(\alpha), \qquad \sigma(f(A)) = \text{ess-range}(f \circ \alpha). \tag{5.18}$$

The second identify implies that

$$\|f(A)\| = \|f \circ \alpha\|_\infty,$$

where the L^∞ norm is defined with respect to μ. Moreover, by the first identity in (5.18), the complement of $\alpha^{-1}(\sigma(A))$ in X has measure zero, implying that

$$\|f \circ \alpha\|_\infty \leq \sup_{\lambda \in \sigma(A)} |f(\lambda)|.$$

This proves the inequality (5.17).

To prove the second claim in (b), suppose that f is continuous. For $\lambda \in \sigma(A)$, (5.18) implies that

$$\mu\big\{x \in X : \lambda - \delta < \alpha(x) < \lambda + \delta\big\} > 0 \tag{5.19}$$

for all $\delta > 0$. Since the preimage under f of an open neighborhood of $f(\lambda)$ is open, (5.19) implies that

$$\mu\big\{x \in X : f(\lambda) - \varepsilon < f \circ \alpha(x) < f(\lambda) + \varepsilon\big\} > 0$$

for $\varepsilon > 0$. Hence, $f(\lambda) \in \sigma_{\text{ess}}(f \circ \alpha)$. This shows that

$$f(\sigma(A)) \subset \sigma(f(A)),$$

which gives

$$\sup_{\lambda \in \sigma(A)} |f(\lambda)| \leq \|f(A)\|$$

when f is continuous.

For (c), under the assumption that $f_n \to f$ pointwise with $\sup |f_n| \leq M$, the dominated convergence theorem implies that

$$\lim_{n \to 0} \big\|(f_n \circ \alpha - f \circ \alpha)v\big\| = 0$$

for each $v \in L^2(X, d\mu)$. By (5.16) this implies that $f_n(A)v \to f(A)v$.

The final step is to establish uniqueness. Let $\mathcal{B}_b(\mathbb{R})$ denote the space of bounded Borel functions, and suppose that Φ_1 and Φ_2 are maps $\mathcal{B}_b(\mathbb{R}) \to \mathcal{L}(\mathcal{H})$ with the properties listed in the theorem. These maps agree on linear combinations of the form $\sum_j c_j h_{z_j}$, with z_j strictly complex, by property (c). The Stone–Weierstrass theorem shows that the span of such linear combinations is dense, with respect to $\|\cdot\|_\infty$, in the space of bounded continuous functions $C_b(\mathbb{R})$. Hence Φ_1 and Φ_2 agree on $C_b(\mathbb{R})$ by (b).

Now set

$$\mathcal{A} := \big\{ f \in \mathcal{B}_b(\mathbb{R}) : \Phi_1(f) = \Phi_2(f) \big\},$$

which is an algebra of functions by (a) and the linearity of the Φ_j. Furthermore, $\mathcal{A}$ is an algebra and closed under pointwise limits of uniformly bounded functions. The characteristic function of a closed interval can be realized as a pointwise limit of bounded continuous functions. Hence $\chi_I \in \mathcal{A}$ for a closed interval $I \subset \mathbb{R}$.

The collection of supports of functions in $\mathcal{A}$ forms a σ-algebra, and since this σ-algebra contains all closed intervals it also contains all Borel sets. Thus $\mathcal{A}$ contains all simple Borel functions. Since every Borel function f is a pointwise limit of simple Borel functions f_n with $|f_n| \leq |f|$, it follows that $\mathcal{A} = \mathcal{B}_b(\mathbb{R})$. □

One immediate consequence of part (b) of Theorem 5.9 is the fact that the resolvent estimate of Corollary 4.11 becomes, for A self-adjoint, the equality

$$\big\|(A - z)^{-1}\big\| = \frac{1}{\operatorname{dist}(z, \sigma(A))}. \tag{5.20}$$

This relation can be quite helpful in locating the spectrum. For $z \in \mathbb{C}$, suppose that we can find a unit vector $u \in \mathcal{D}(A)$ such that

$$\|(A - z)u\| \leq \varepsilon.$$

This implies that $\|(A - z)^{-1}\| \geq 1/\varepsilon$, and therefore by (4.27) there exists a point $\lambda \in \sigma(A)$ with $|\lambda - z| \leq \varepsilon$.

One potential problem with the construction of the functional calculus given here is the fact that measure space (X, μ) provided by the spectral theorem is not specified explicitly (and not even uniquely defined). It is therefore very useful to express the functional calculus in terms of the resolvent, without reference to the auxiliary measure space.

Theorem 5.10. *Suppose A is self-adjoint and $f \in C_b(\mathbb{R})$. Then*

$$f(A) = \lim_{\varepsilon \to 0} \frac{1}{2\pi i} \int_{-\infty}^{\infty} f(\lambda) \Big[(A - \lambda - i\varepsilon)^{-1} - (A - \lambda + i\varepsilon)^{-1} \Big] \, d\lambda, \tag{5.21}$$

with the limit taken in the strong operator sense. If f is uniformly continuous, then the limit $\varepsilon \to 0$ exists in the operator-norm topology.

Proof By the functional calculus, the operator on the right-hand side of (5.21) is given by $f_\varepsilon(A)$, where

$$\begin{aligned} f_\varepsilon(x) &:= \frac{1}{2\pi i}\int_{-\infty}^{\infty} f(\lambda)\left[\frac{1}{x-\lambda-i\varepsilon} - \frac{1}{x-\lambda+i\varepsilon}\right] d\lambda \\ &= \frac{\varepsilon}{\pi}\int_{-\infty}^{\infty} \frac{f(\lambda)}{(x-\lambda)^2+\varepsilon^2}\, d\lambda. \end{aligned}$$

Subtracting this from $f(x)$ and making the change of variables $x = \lambda - \varepsilon t$ give

$$f(x) - f_\varepsilon(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{f(x) - f(x+\varepsilon t)}{t^2+1}\, dt \tag{5.22}$$

Since f is continuous, $f_\varepsilon \to f$ pointwise as $\varepsilon \to 0$, by the dominated convergence theorem. Furthermore, $\|f_\varepsilon\|_\infty \le \|f\|_\infty$ so the strong operator limit in (5.21) follows from Theorem 5.9(c).

If f is uniformly continuous, then it is clear from (5.22) that $f_\varepsilon \to f$ uniformly. In this case, operator-norm convergence follows from Theorem 5.9(b). □

5.4 Spectral Decomposition

In this section we will discuss various decompositions of the spectrum of a self-adjoint operator A. We have already defined one important subset, the point spectrum $\sigma_{\mathrm{pt}}(A)$ consisting of the eigenvalues.

To identify other components of the spectrum, we can use the functional calculus to create projections. To each Borel subset $E \subset \mathbb{R}$, we associate an orthogonal projection

$$\Pi_E := \chi_E(A),$$

as defined by Theorem 5.9. This yields a family of projections, collectively denoted by Π, called the *spectral resolution* of A. The following result shows the relationship between the spectrum and the support of Π.

Theorem 5.11. *For a self-adjoint operator A, the point $\lambda \in \mathbb{R}$ lies in $\sigma(A)$ if and only if $\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)} \neq 0$ for all $\varepsilon > 0$. If $\Pi_{\{\lambda\}} \neq 0$, then λ is an eigenvalue and the range of $\Pi_{\{\lambda\}}$ is the corresponding eigenspace.*

Proof Define the measure space (X, μ), the unitary map $Q : L^2(X, d\mu) \to \mathcal{H}$, and $\alpha(x) := x$ as in Theorem 5.6. For $E \subset \mathbb{R}$,

$$\Pi_E = Q^{-1}\chi_{\alpha^{-1}\{E\}}Q.$$

Therefore $\Pi_E \neq 0$ if and only if $\mu\{\alpha \in E\} > 0$, and the first claim follows from the fact that $\sigma(A)$ is the essential range of α, by Theorem 4.5.

If $\phi \in \text{range}(\Pi_{\{\lambda\}})$, then $Q^{-1}\phi$ has support on $\alpha = \lambda$. Thus $\alpha Q^{-1}\phi = \lambda Q^{-1}\phi$, which implies that $A\phi = \lambda\phi$. □

5.4.1 Discrete and Essential Spectrum

Our first subdivision of the spectrum takes into account the rank of the spectral projection Π near a point.

Definition 5.12. For a self-adjoint operator A, the *essential spectrum* $\sigma_{\text{ess}}(A)$ is the set of $\lambda \in \sigma(A)$ such that $\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}$ has infinite rank for all $\varepsilon > 0$. The *discrete spectrum* $\sigma_{\text{disc}}(A)$ consists of $\lambda \in \sigma(A)$ for which $\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}$ has finite rank for some $\varepsilon > 0$.

Note that the definitions are complementary, so that

$$\sigma(A) = \sigma_{\text{disc}}(A) \cup \sigma_{\text{ess}}(A) \tag{5.23}$$

is a disjoint union. The discrete spectrum is clearly a subset of the point spectrum. The difference between them is that an eigenvalue may have infinite multiplicity, or otherwise lie within the essential spectrum. Such eigenvalues are considered point but not discrete spectrum.

We saw in Theorem 4.16 that elements of the spectrum are approximate eigenvalues in the sense that there exists a sequence of unit vectors u_n such that $(A - \lambda)u_n \to 0$. We can describe the essential spectrum in a similar way by adding an extra requirement that $u_n \to 0$ in the weak sense. Note that this condition is satisfied in particular if $\{u_n\}$ is orthonormal, by Exercise 2.7.

Theorem 5.13 (Weyl's Criterion). *Suppose A is a self-adjoint on a Hilbert space $\mathcal{H}$. A point $z \in \mathbb{C}$ lies in $\sigma_{\text{ess}}(A)$ if and only if there exists a sequence $\{u_n\} \subset \mathcal{D}(A)$ with $\|u_n\| = 1$, such that $u_n \to 0$ in the weak sense and*

$$\lim_{n\to\infty} \|(A - z)u_n\| = 0.$$

Proof We already know from Theorem 4.16 that no such sequence exists for $z \in \rho(A)$, so it suffices to consider $\lambda \in \sigma(A)$. If $\lambda \in \sigma_{\text{ess}}(A)$, then range $\Pi_{(\lambda-\frac{1}{n},\lambda+\frac{1}{n})}$ is infinite-dimensional for all n. Hence, for each n we can choose a unit vector $u_n \in \text{range}\,\Pi_{(\lambda-\frac{1}{n},\lambda+\frac{1}{n})}$ which is orthogonal to $u_1, \ldots, u_{n-1}$. An orthonormal sequence converges weakly to 0, by Exercise 2.7. Furthermore, the fact that $u_n \in \text{range}\,\Pi_{(\lambda-\frac{1}{n},\lambda+\frac{1}{n})}$ implies

$$\|(A - \lambda)u_n\| < \frac{1}{n}.$$

Hence $(A - \lambda)u_n \to 0$.

Now suppose $\lambda \in \sigma_{\text{disc}}(A)$ and choose $\varepsilon > 0$ so that $\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}$ has finite rank. Let $\{e_1, \ldots, e_k\}$ be a basis for range $\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}$, so that

$$\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}u = \sum_{j=1}^{k} \langle u, e_j \rangle e_j.$$

Let $\{u_n\}$ be a sequence in $\mathcal{D}(A)$ with $\|u_n\| = 1$. If $u_n \to 0$ weakly, then in particular

$$\lim_{n\to\infty} \langle u_n, e_j \rangle = 0$$

for all $j \in \mathbb{N}$. Hence

$$\lim_{n\to\infty} \|\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}u_n\| = 0. \tag{5.24}$$

Since A commutes with Π_E we can use orthogonal decomposition to estimate

$$\begin{aligned}\|(A-\lambda)u_n\| &= \left\|(A-\lambda)(1-\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)})u_n\right\| + \left\|(A-\lambda)\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}u_n\right\| \\ &\geq \varepsilon\left\|(1-\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)})u_n\right\| - \varepsilon\left\|\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}u_n\right\|.\end{aligned}$$

By (5.24), this implies that

$$\liminf_{n\to\infty} \|(A-\lambda)u_n\| \geq \varepsilon.$$

Therefore, the criteria for $\{u_n\}$ cannot be satisfied for $\lambda \in \sigma_{\text{disc}}(A)$. □

The sequence $\{u_n\}$ appearing in Theorem 5.13 is called a *Weyl sequence* for λ. If A is the closure of an essentially self-adjoint operator A_0, then it suffices to consider Weyl sequences in $\mathcal{D}(A_0)$. In other words, it suffices to consider functions in a core domain for A. To see this, suppose $\{u_n\} \subset \mathcal{D}(A)$ is a Weyl sequence for λ. Because A is the closure of A_0, for each n, we can find $w_n \in \mathcal{D}(A_0)$ such that $\|w_n - u_n\| \leq 1/n$ and $\|Aw_n - Au_n\| \leq 1/n$. The sequence $\{w_n/\|w_n\|\}$ then gives a Weyl sequence for λ contained in $\mathcal{D}(A_0)$.

The term "essential spectrum" is explained by the following corollary to Theorem 5.13. Since compact operators map weakly convergent sequences to strongly convergent by Theorem 3.38, the addition of a compact operator has no effect on the existence of Weyl sequences. This yields the proof of the following:

Theorem 5.14 (Weyl Stability). *Suppose that A and B are self-adjoint operators, with B compact. Then*

$$\sigma_{\text{ess}}(A+B) = \sigma_{\text{ess}}(A).$$

5.4.2 Continuous Spectrum

Another way to classify the spectrum is through the properties of the associated measures. By the Lebesgue decomposition theorem, a regular Borel measure μ on $\mathbb{R}$ admits a unique decomposition of the form,

$$\mu = \mu_{\mathrm{pp}} + \mu_{\mathrm{ac}} + \mu_{\mathrm{sc}}. \tag{5.25}$$

Here μ_{pp} is a *pure point* measure (a sum of point measures), μ_{ac} is absolutely continuous with respect to Lebesgue measure, and μ_{sc} is singular continuous. See Appendix A.1.5 for a brief review of this material.

Since the measure space (X, μ) provided by the spectral theorem consists of copies of $\mathbb{R}$ equipped with finite Borel measures, we can derive from (5.25) the decomposition,

$$L^2(X, d\mu) = L^2(X, d\mu_{\mathrm{pp}}) \oplus L^2(X, d\mu_{\mathrm{ac}}) \oplus L^2(X, d\mu_{\mathrm{sc}}).$$

The corresponding decomposition of $\mathcal{H}$ is denoted

$$\mathcal{H} = \mathcal{H}_{\mathrm{pp}} \oplus \mathcal{H}_{\mathrm{ac}} \oplus \mathcal{H}_{\mathrm{sc}}.$$

Based on this subdivision, the continuous portion of the spectrum can be decomposed in disjoint sets,

$$\sigma_{\mathrm{cont}}(A) := \sigma_{\mathrm{ac}}(A) \cup \sigma_{\mathrm{sc}}(A),$$

where

$$\sigma_{\mathrm{ac}}(A) := \sigma(A|_{\mathcal{H}_{\mathrm{ac}}}), \qquad \sigma_{\mathrm{sc}}(A) := \sigma(A|_{\mathcal{H}_{\mathrm{sc}}}).$$

Note that the point spectrum $\sigma_{\mathrm{pt}}(A)$, the set of eigenvalues, is not necessarily closed. On the other hand, the spectrum of the restriction $A|_{\mathcal{H}_{\mathrm{pp}}}$ is closed by definition. In fact, it is easy to check that

$$\sigma(A|_{\mathcal{H}_{\mathrm{pp}}}) = \overline{\sigma_{\mathrm{pt}}(A)}.$$

(This follows from Exercise 4.9.) Hence the measure decomposition (5.25) leads to the partition,

$$\sigma(A) = \overline{\sigma_{\mathrm{pt}}(A)} \cup \sigma_{\mathrm{ac}}(A) \cup \sigma_{\mathrm{sc}}(A), \tag{5.26}$$

as an alternative to (5.23).

Unlike the essential spectrum, the decomposition (5.26) is unstable under compact perturbations. A remarkable theorem of Weyl and von Neumann [93] says

that for any bounded self-adjoint operator, there exists a compact (in fact, Hilbert–Schmidt) perturbation such that the perturbed operator has pure point spectrum. On the other hand, Kato [51] proved that the absolutely continuous spectrum is stable under trace-class perturbations.

For $v \in \mathcal{H}$, the spectral resolution can be used to associate a *spectral measure* to a vector $v \in \mathcal{H}$, by setting

$$\mu_v(E) := \langle v, \Pi_E v \rangle$$

for a Borel set $E \subset \mathbb{R}$. Integration with respect to the spectral measure is related to the functional calculus by

$$\langle v, f(A)v \rangle = \int_{\mathbb{R}} f \, d\mu_v \tag{5.27}$$

for $f \in \mathcal{B}_b(\mathbb{R})$. This is analogous to the formula (5.10) for the spectral measure in the unitary case.

Starting from the quadratic functional (5.27), we can derive the corresponding sesquilinear form $h(v, w) = \langle v, f(A)w \rangle$ using the polarization identity as in (2.17). By the Riesz lemma, the operator $f(A)$ is uniquely determined by the set of matrix elements $\langle v, f(A)v \rangle$. Therefore, the full functional calculus could be recovered from knowledge of the spectral resolution Π.

At the operator level, the relationship (5.27) can be stated more directly in terms of integration with respect to the *projection-valued measure* $E \mapsto \Pi_E$. That is, it is possible to define integration with respect to $d\Pi$, just as for an ordinary measure. In terms of this operator integral, (5.27) is equivalent to

$$f(A) = \int_{\mathbb{R}} f(\lambda) \, d\Pi(\lambda). \tag{5.28}$$

We will not develop the projection-valued measure theory here. It suffices for our purposes to regard (5.28) as a shorthand notation for the weak definition (5.27).

5.4.3 The Min–Max Principle

The Weyl criterion of Theorem 5.13 allows us to locate the essential spectrum without explicit knowledge of the resolvent or spectral projectors. In that sense, it is similar to the max–min principle which we stated for compact operators in Theorem 4.22. For general self-adjoint operators, we can formulate a version of this principle which helps to separate the bottom of the essential spectrum from the discrete eigenvalues below it.

Theorem 5.15 (Min–Max Principle). *Let A be a self-adjoint operator whose spectrum is bounded below. Let Λ_k denote the set of subspaces of $\mathcal{D}(A)$ of dimension k, and define*

$$\alpha_k := \min_{W \in \Lambda_k} \left\{ \max_{u \in W \setminus \{0\}} \frac{\langle u, Au \rangle}{\|u\|^2} \right\} \tag{5.29}$$

for $k \in \mathbb{N}$. Then for each k, one of the following alternatives holds:

(a) α_k *is the kth eigenvalue (arranged in increasing order and counted with multiplicity) and there are at least k eigenvalues below the essential spectrum.*
(b) $\alpha_k = \inf \sigma_{\mathrm{ess}}(A)$ *and there are at most $k - 1$ eigenvalues below the essential spectrum.*

Proof For $c \in \mathbb{R}$, suppose that $\operatorname{rank} \Pi_{(-\infty,c)} \geq k$. Since the spectrum of A is bounded below, $\operatorname{range} \Pi_{(-\infty,c)} \subset \operatorname{range} \Pi_{(a,c)}$ for some $a \in \mathbb{R}$. By the characterization of $\mathcal{D}(A)$ in Theorem 5.6, it follows that

$$\operatorname{range} \Pi_{(-\infty,c)} \subset \mathcal{D}(A).$$

The assumption on the rank thus implies that there is a subspace $W \in \Lambda_k$ for which $W \subset \operatorname{range} \Pi_{(-\infty,c)}$. The restriction of A to $\operatorname{range} \Pi_{(-\infty,c)}$ is bounded by c, so that

$$\max_{u \in W \setminus \{0\}} \frac{\langle u, Au \rangle}{\|u\|^2} \leq c.$$

Therefore,

$$\alpha_k \leq c.$$

Now suppose that $\operatorname{rank} \Pi_{(-\infty,c)} < k$. Then, for each subspace $W \in \Lambda_k$, there exists some vector $u \in W$ such that

$$\langle u, Au \rangle > c\|u\|^2.$$

This implies that

$$\alpha_k \geq c.$$

Taking the contrapositive of these statements, we conclude that for all $\varepsilon > 0$,

$$\begin{aligned} \operatorname{rank} \Pi_{(-\infty,\alpha_k-\varepsilon)} &< k, \\ \operatorname{rank} \Pi_{(-\infty,\alpha_k+\varepsilon)} &\geq k. \end{aligned} \tag{5.30}$$

It follows that $\operatorname{rank} \Pi_{(\alpha_k-\varepsilon,\alpha_k+\varepsilon)} \geq 1$ for all $\varepsilon > 0$, which means that $\alpha_k \in \sigma(A)$. If $\Pi_{(\alpha_k-\varepsilon,\alpha_k+\varepsilon)}$ has finite rank for some $\varepsilon > 0$, then $\alpha_k \in \sigma_{\mathrm{disc}}(A)$, and it then follows from (5.30) that α_k is the kth eigenvalue.

On the other hand, if $\Pi_{(\alpha_k-\varepsilon,\alpha_k+\varepsilon)}$ has infinite rank for some $\varepsilon > 0$, then $\alpha_k \in \sigma_{\mathrm{ess}}(A)$, by definition. The first statement of (5.30) implies that no point below α_k lies in $\sigma_{\mathrm{ess}}(A)$, so α_k is the bottom of the essential spectrum. □

There is a corresponding max–min principle. That is, the value of α_k from Theorem 5.15 can be computed as

$$\alpha_k = \max_{v_1,\ldots,v_{k-1}\in\mathcal{H}} \left\{ \min_{u\in\mathcal{D}(A)\cap\{v_1,\ldots,v_{k-1}\}^\perp\setminus\{0\}} \frac{\langle u, Au\rangle}{\|u\|^2} \right\}. \tag{5.31}$$

We leave the proof to Exercise 5.8.

5.5 Exercises

5.1. For a self-adjoint operator A, let U be the Cayley transform defined in Lemma 5.7. Prove that $1 \in \rho(U)$ if and only if A is bounded.

5.2. Let A be a self-adjoint operator.

(a) Suppose that λ is an isolated point of $\sigma(A)$, meaning that there is a neighborhood of λ containing no other point of $\sigma(A)$. Prove that λ is an eigenvalue of A.
(b) Prove that if the spectrum of A is purely discrete, in the sense of Definition 5.12, if and only if A has compact resolvent (as defined in Exercise 4.11).

5.3. If $\{A_n\}$ and A are self-adjoint operators, then we say that $A_n \to A$ in the *norm-resolvent* sense if

$$\lim_{n\to\infty} \left\| (A_n - i)^{-1} - (A - i)^{-1} \right\| = 0. \tag{5.32}$$

(a) Prove that norm-resolvent convergence implies that

$$\lim_{n\to\infty} \left\| (A_n - z)^{-1} - (A - z)^{-1} \right\| = 0,$$

for all z with $\operatorname{Im} z \neq 0$.
(b) For $x \in \rho(A) \cap \mathbb{R}$, show that $x \in \rho(A_n)$ for all n sufficiently large and that $(A_n - x)^{-1} \to (A - x)^{-1}$ in the operator topology.

5.4. Assume that $\{A_n\}$ and A are self-adjoint operators with $A_n \to A$ in the norm-resolvent sense, as defined in Exercise 5.3. Suppose f is a bounded continuous function on $\mathbb{R}$ whose limits at $\pm\infty$ exist and are equal. Prove that

$$\lim_{n\to\infty} \left\| f(A_n) - f(A) \right\| = 0.$$

[Hint: Approximate $f(A_n)$ and $f(A)$ by polynomials in the respective Cayley transforms U_n and U.]

5.5. For a self-adjoint operator A and $t \in \mathbb{R}$, let

$$U(t) := e^{itA},$$

as defined by the functional calculus. It follows from Theorem 5.9 that $\{U(t)\}_{t\in\mathbb{R}}$ is a *unitary group*, meaning that each $U(t)$ is unitary and

$$U(s+t) = U(s)U(t)$$

for all $s, t \in \mathbb{R}$. For $v \in \mathcal{D}(A)$, prove that

$$\lim_{t\to 0} \frac{1}{t}\big[U(t)v - v\big] = iAv.$$

5.6. Let $U(t) := e^{itA}$ be the unitary group associated with a self-adjoint operator A as in Exercise 5.5. Note that $t \mapsto U(t)$ is continuous in the strong operator topology by Theorem 5.9. Prove that $U(t)$ is continuous with respect to the operator-norm topology if and only if A is bounded.

5.7. Suppose A is self-adjoint, and let γ be a simple, positively oriented piecewise smooth curve in $\rho(A)$ that encloses a subset $E \subset \sigma(A)$. Prove that the spectral projection onto E is given by

$$\Pi_E = \frac{-1}{2\pi i} \int_\gamma (A - z)^{-1}\, dz,$$

where the contour integral is defined as in Lemma 4.7.

5.8. Prove the max–min principle (5.31).

5.9. Suppose that A is a self-adjoint operator on a Hilbert space $\mathcal{H}$. Prove that the interval $(\lambda - \varepsilon, \lambda + \varepsilon)$ intersects the essential spectrum of A if and only if there exists an infinite-dimensional subspace $W \subset \mathcal{D}(A)$ such that

$$\|(A-\lambda)w\| \le \varepsilon\|w\|$$

for all $w \in W$.

Notes

The strategy of proving the spectral theorem for unitary operators using Fourier series was inspired by a set of lecture notes by Michael Taylor. A similar method is presented in Simon [83, §5.5].

Other expositions of the von Neumann trick of deriving the spectral theorem for unbounded operators from the Cayley transform can be found in Simon [83, §7.2] and Taylor [90, §8.1]. In Weidmann [94, §7.3], the spectral theorem for unbounded operators is proven by using the resolvent formula (5.21) to derive the spectral resolution.

Another way to prove the spectral theorem for unitary operators is to construct the continuous functional calculus for bounded self-adjoint operators directly, as described in Section 5.2. This construction, which extends easily to normal operators, is developed for example in Reed and Simon [69, Chapter VII] and Hall [41, Chapter 10]. An alternate route to the spectral theorem for normal operators is via Banach algebra theory, as in MacCluer [60, §6.1] or Rudin [79, Part III].

One-parameter groups generated by self-adjoint operators, such as the unitary group introduced in Exercise 5.5, are an important topic in spectral theory, particular in applications to quantum mechanics. For more details, see Reed and Simon [69, §VIII.4], Schmüdgen [80, Chapter 6], or Weidmann [94, §7.6].

Chapter 6
The Laplacian with Boundary Conditions



The oldest problem in spectral theory is to understand the sound of a vibrating string. As we noted in Chapter 1 this problem dates back to Pythagoras. The first mathematical model for the string was the one-dimensional wave equation developed by Jean d'Alembert in 1746. If $u(x, t)$ denotes the displacement of the string at position $x \in [0, \ell]$ and time t, the equation takes the form,

$$(\partial_t^2 - \partial_x^2)u = 0 \tag{6.1}$$

(with physical constants omitted). Separating the time and position variables yields a spatial equation,

$$-\Delta\phi = \lambda\phi, \quad u(0) = u(\ell) = 0. \tag{6.2}$$

Historically this is called the Helmholtz equation, based on Hermann von Helmholtz's work on electrodynamics in the late nineteenth century. In the later terminology of Hilbert we would describe it as the "eigenvalue equation" for the Laplacian with Dirichlet boundary conditions.

Equation (6.2) has an obvious family of solutions,

$$\phi_k(x) := \sin\left(\frac{k\pi x}{\ell}\right)$$

for $k \in \mathbb{N}$, with eigenvalues

$$\lambda_k = \left(\frac{k\pi}{\ell}\right)^2.$$

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1_6

If we reinstate the physical constants in (6.1), then this model predicts a set of vibrational frequencies,

$$\nu_k = \frac{k}{2\ell}\sqrt{\frac{T}{\rho}}, \qquad k \in \mathbb{N}, \tag{6.3}$$

measured in Hz (cycles per second), where ρ is the linear density of the string and T the tension. The formula (6.3) reproduces an empirical law for the overtone series discovered by Mersenne.

In this chapter, we will analyze higher dimensional analogues of the string problem. That is, we consider the spectral theory of the Laplacian on an open set $\Omega \in \mathbb{R}^n$. We will concentrate on the classical boundary conditions, Dirichlet and Neumann, and develop the corresponding self-adjoint extensions.

In the Dirichlet case, for Ω bounded we will show that the spectrum of $\sigma(-\Delta)$ is purely discrete. In this case the Dirichlet Laplacian admits an orthonormal basis of eigenfunctions, with a sequence of eigenvalues

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \to +\infty. \tag{6.4}$$

A similar result holds in the Neumann case, provided the boundary $\partial\Omega$ is sufficiently regular, except that $\lambda_1 = 0$.

In dimension two, the sequence of Dirichlet eigenvalues yields the set of vibrational frequencies for a membrane which is fixed at its edges, i.e., the overtone series for a drum with a head given by Ω. As in the string example, the eigenvalue equation arises from a simplification of the physical model, a linear approximation that ignores friction.

Example 6.1. Consider the Laplacian with Dirichlet boundary conditions on an open rectangle $\Omega = (0, \ell_1) \times \cdots \times (0, \ell_n) \subset \mathbb{R}^n$. Separation of variables reduces the eigenvalue equation to a system of one-dimensional equations. The resulting eigenfunctions are

$$\phi_k(x) := \prod_{j=1}^{n} \frac{1}{\sqrt{\ell_j}} \sin\left(\frac{\pi k_j x_j}{\ell_j}\right) \tag{6.5}$$

for $k \in \mathbb{N}^n$. Some two-dimensional examples are illustrated in Figure 6.1.

We can show that $\{\phi_k\}_{k\in\mathbb{N}^n}$ forms a basis for $L^2(\Omega)$ by exploiting the obvious connection to Fourier series. For convenience, let us first change variables so that $\ell_j = \pi$ for each j. We can identify $L^2((0,\pi)^n)$ with a closed subspace of $L^2(\mathbb{T}^n)$, where $\mathbb{T}^n := \mathbb{R}^n/(2\pi\mathbb{Z})^n$, by first extending functions on $(0,\pi)^n$ to odd functions on $(-\pi,\pi)^n$ and then making them 2π-periodic. These extended functions can then be expanded in terms of the Fourier basis for $L^2(\mathbb{T}^n)$ (see Example 2.32). For odd functions in particular, the expansion reduces to a Fourier sine series. It follows

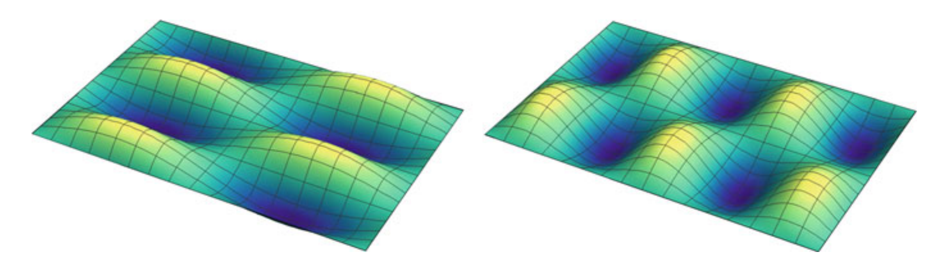


Fig. 6.1 Dirichlet eigenfunctions for a rectangular domain

that the functions (6.5) yield an orthonormal basis for $L^2((0,\pi)^n)$. For the original rectangle Ω this shows that the full Dirichlet spectrum is

$$\sigma(-\Delta) = \left\{ \sum_{j=1}^{n} \frac{\pi^2 k_j^2}{\ell_j^2} \,:\, k \in \mathbb{N}^n \right\}.$$

Under Neumann conditions, the only change is to switch from sines to cosines in (6.5). For the Fourier basis argument we then use even extensions instead of odd. The set of eigenvalues is given by the same formula, except that each k_j is allowed to be zero. ♢

Example 6.2. Consider the Laplacian on the unit disk $\mathbb{D} \subset \mathbb{R}^2$. In polar coordinates,

$$\Delta = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}.$$

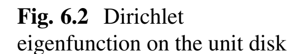
If we substitute $\phi(r,\theta) = h(r)e^{ik\theta}$ into the eigenvalue equation $-\Delta\phi = \lambda\phi$, then the equation for the radial factor is

$$\left(r\frac{\partial}{\partial r}\right)^2 h + (\lambda r^2 - k^2)h = 0.$$

The solutions which are regular as $r \to 0$ are given by $h(r) = J_k(\sqrt{\lambda}r)$, where J_k is the standard Bessel function. To satisfy $h(1) = 0$, we set $\sqrt{\lambda} = j_{k,m}$, where $\{0 < j_{k,1} < j_{k,2} < \dots\}$ denotes the sequence of zeros of J_k. This gives a set of eigenfunctions

$$\phi_{k,m}(r,\theta) = J_k(j_{k,m}r)e^{ik\theta}.$$

An example is shown in Figure 6.2.



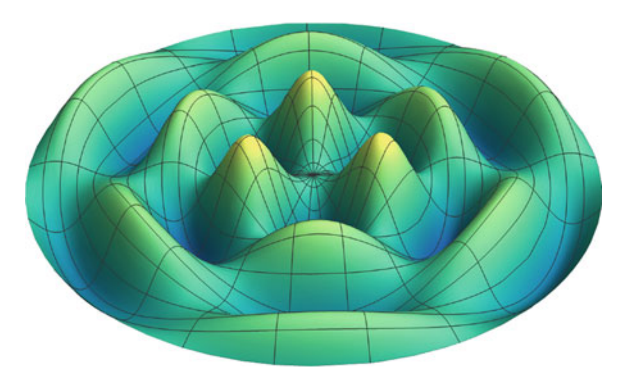


Fig. 6.2 Dirichlet eigenfunction on the unit disk

This set of eigenfunctions yields a basis for $L^2(\mathbb{D})$, so that

$$\sigma(-\Delta) = \left\{ j_{k,m}^2 \,:\, k \in \mathbb{Z}, m \in \mathbb{N} \right\}.$$

To prove this, one can use Fubini's theorem and the Fourier basis theorem to reduce the argument to the fact that $\{\sqrt{r}\, J_k(j_{k,m} r)\}_{k\in\mathbb{Z},m\in\mathbb{N}}$ is a basis for $L^2(0, 1)$. This result is well known from special function theory, although the proof is not exactly elementary.

The same derivation applies in the case of Neumann boundary conditions, except that the zeros of $J_k(z)$ are replaced with the zeros of $J_k'(z)$, which are denoted by $j_{k,m}'$. Figure 6.3 shows an example of a Neumann eigenfunction. ◊

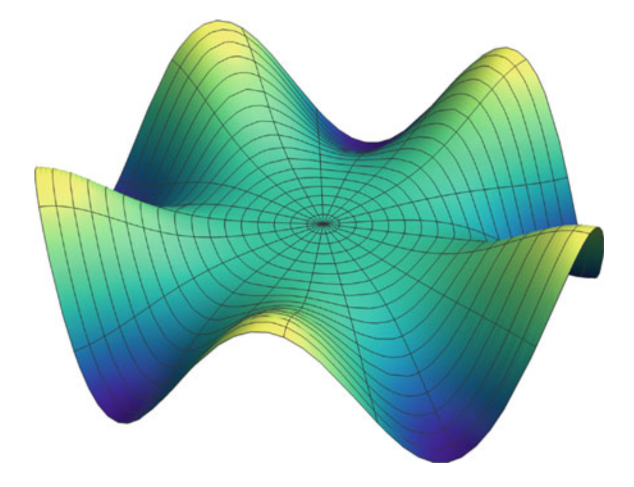


Fig. 6.3 Neumann eigenfunction on the unit disk

6.1 Self-adjoint Extensions

As noted in Example 3.20, for a domain $\Omega \subset \mathbb{R}^n$ we can impose Dirichlet or Neumann boundary conditions to guarantee positivity of the Laplacian, provided there is sufficient regularity to justify the use of Green's identity. In this section, we will produce self-adjoint extensions of $-\Delta$ corresponding to these classical boundary conditions. These are Friedrichs extensions, constructed using the quadratic form approach developed in Section 3.4.3. In both cases, self-adjointness does not require any restriction on the regularity of $\partial\Omega$.

Although we focus on the Laplacian to simplify the exposition, the methods discussed here extend directly to the case of uniformly elliptic operators. (See Exercise 6.4.)

6.1.1 The Space $H_0^1(\Omega)$

Let $\Omega \subset \mathbb{R}^n$ be an open set. Because functions in $L^2(\Omega)$ are defined only up to sets of measure zero, they do not have well-defined boundary values on $\partial\Omega$. However, functions in the Sobolev space $H^1(\Omega)$ defined in Section 2.5.2 have enough regularity to allow for a meaningful interpretation of the Dirichlet conditions $u|_{\partial\Omega} = 0$. To implement these conditions, we define the subspace

$$H_0^1(\Omega) := \overline{C_0^\infty(\Omega)} \subset H^1(\Omega), \tag{6.6}$$

where the closure is with respect to the H^1 norm. As a closed subspace, $H_0^1(\Omega)$ is itself a Hilbert space.

It is not immediately clear that $H_0^1(\Omega)$ differs from $H^1(\Omega)$. After all, the closure of $C_0^\infty(\Omega)$ with respect to the L^2 norm is simply $L^2(\Omega)$. Before we proceed, let us consider how the definition of H_0^1 works in some special cases.

Example 6.3. Let Ω be the unit interval $(0, 1)$. In Example 2.24 we saw that functions in $H^1(0, 1)$ are absolutely continuous. We claim that $H_0^1(0, 1)$ is the corresponding space with classical Dirichlet conditions at the endpoints, i.e.,

$$H_0^1(0, 1) = \left\{ f \in AC[0, 1] : \ f' \in L^2[0, 1], \ f(0) = f(1) = 0 \right\}.$$

The endpoint condition follows from the fact that convergence in H^1 implies uniform convergence in dimension one. To prove this, we need to estimate $\|f\|_\infty$ in terms of $\|f\|_{H^1}$. For a function $f \in H^1(0, 1)$, there exists a point $x_1 \in [0, 1]$ at which

$$|f(x_1)| = \|f\|_\infty,$$

by continuity. For $x \in [0, 1]$, we can use the formula

$$f(x_1) = f(x) + \int_x^{x_1} f'(t)\, dt \tag{6.7}$$

and apply Cauchy–Schwarz to the integral to obtain

$$\|f\|_\infty \le |f(x)| + |x - x_1|^{\frac{1}{2}} \|f'\|.$$

By squaring and integrating this estimate over x, we can deduce that

$$\|f\|_\infty \le C\|f\|_{H^1}. \tag{6.8}$$

For $f \in H_0^1(0, 1)$, there exists a sequence $\{\psi_k\} \subset C_0^\infty(0, 1)$ which converges to f in the H^1 norm. By (6.8) this convergence is uniform, and so f must vanish at the endpoints. ♢

Example 6.4. Suppose that $\Omega \subset \mathbb{R}^n$ has a piecewise C^1 boundary, and that $u \in C^\infty(\overline{\Omega})$. We will show that if u also lies in $H_0^1(\Omega)$, then it satisfies classical Dirichlet boundary conditions.

Consider an approximating sequence $\{\psi_n\} \subset C_0^\infty(\Omega)$ such that $\psi_n \to u$ with respect to the H^1 norm. For $u, \phi \in C^\infty(\overline{\Omega})$, Green's first identity implies that

$$\begin{aligned}
\int_{\partial\Omega} u \frac{\partial \phi}{\partial \nu}\, dS &= \int_\Omega \big[\nabla u \cdot \nabla \phi + u \Delta \phi\big] dx \\
&= \lim_{n\to\infty} \int_\Omega \big[\nabla \psi_n \cdot \nabla \phi + \psi_n \Delta \phi\big] dx \\
&= 0.
\end{aligned}$$

Since this holds for all $\phi \in C^\infty(\overline{\Omega})$, it follows that

$$u|_{\partial\Omega} = 0.$$

This conclusion can be generalized to all of $H_0^1(\Omega)$ through the theory of boundary "traces" of Sobolev functions; see, for example, Evans [29, §5.5]. ♢

If Ω is bounded, then functions in $H_0^1(\Omega)$ satisfy a classical inequality that is closely connected to the spectral theory of the Dirichlet Laplacian on Ω.

Theorem 6.5 (Poincaré Inequality). *Suppose $\Omega \subset \mathbb{R}^n$ is a bounded open set. There exists a constant $\gamma > 0$ such that*

$$\|u\|_{L^2} \le \gamma \|\nabla u\|_{L^2}$$

for all $u \in H_0^1(\Omega)$.

Proof By the definition of $H_0^1(\Omega)$, it suffices to prove the result for functions in $C_0^\infty(\Omega)$. Fix $M > 0$ so that

$$\Omega \subset \mathcal{R} := [-M, M]^n.$$

The natural embedding $C_0^\infty(\Omega) \subset C_0^\infty(\mathcal{R})$, given by extending by zero, is an isometry with respect to both H^1 and L^2 norms. Thus it suffices to derive the Poincaré inequality for $\psi \in C_0^\infty(\mathcal{R})$.

By the fundamental theorem of calculus,

$$\psi(x) = \int_{-M}^{x_1} \partial_1 \psi(y, x_2, \ldots, x_n)\, dy.$$

Applying the Cauchy–Schwarz inequality on $L^2(-M, M)$ gives the estimate,

$$|\psi(x)|^2 \leq 2M \int_{-M}^{M} \left|\partial_1 \psi(y, x_2, \ldots, x_n)\right|^2 dy,$$

for all $x \in [-M, M]^n$. Integrating this estimate over x then yields

$$\begin{aligned} \|\psi\|^2 &\leq 4M^2 \|\partial_1 \psi\|^2 \\ &\leq 4M^2 \|\nabla \psi\|^2. \end{aligned}$$

□

6.1.2 The Dirichlet Laplacian

Our goal in this section is to obtain a self-adjoint extension of $-\Delta$ on Ω by applying the Friedrichs construction from Section 3.4.3 to the H^1 inner product on $H_0^1(\Omega)$. This case demonstrates one of the primary advantages of the quadratic form approach, namely that the domain of the quadratic form is simpler than the domain of the operator.

Since we are considering multiple extensions in this chapter, let us denote the *Dirichlet Laplacian* by $-\Delta_{\mathrm{D}}$. This operator acts on the domain

$$\mathcal{D}(-\Delta_{\mathrm{D}}) := \left\{ u \in H_0^1(\Omega) : \; -\Delta u \in L^2(\Omega) \right\}, \tag{6.9}$$

where $-\Delta u$ is interpreted in the weak sense. The condition that $-\Delta u \in L^2(\Omega)$ means precisely that there exists $g \in L^2(\Omega)$ such that

$$\langle g, \psi \rangle = \langle u, -\Delta \psi \rangle, \quad \text{for all } \psi \in C_0^\infty(\Omega). \tag{6.10}$$

For $u \in \mathcal{D}(-\Delta_{\mathrm{D}})$, we define $-\Delta u := g$.

Under certain conditions, the definition (6.9) can be simplified. A standard *elliptic regularity* result (Theorem A.20) says that if $\partial\Omega$ is C^2, $u \in H_0^1(\Omega)$, and $-\Delta u \in L^2(\Omega)$, then $u \in H^2(\Omega)$. Thus, in the case of C^2 boundary, the exact domain is

$$\mathcal{D}(-\Delta_{\mathrm{D}}) = H_0^1(\Omega) \cap H^2(\Omega). \tag{6.11}$$

We will discuss the issue of elliptic regularity in more detail in Sections 6.3 and 9.4.2.

To set up the proof of self-adjointness, it will be helpful to rewrite the definition (6.9) in terms of the H^1 inner product. By Green's identity,

$$\langle \nabla\phi, \nabla\psi \rangle = \langle \phi, -\Delta\psi \rangle, \tag{6.12}$$

for $\phi, \psi \in C_0^\infty(\Omega)$. Thus $-\Delta$ is symmetric and positive on $C_0^\infty(\Omega)$.

Taking an approximating sequence $\{\phi_k\} \subset C_0^\infty(\Omega)$ converging to $u \in H_0^1(\Omega)$ thus gives

$$\langle u, \psi \rangle_{H^1} = \langle u, -\Delta\psi \rangle + \langle u, \psi \rangle. \tag{6.13}$$

Therefore, the condition (6.10) is equivalent to the existence of $f \in L^2(\Omega)$ such that

$$\langle u, \psi \rangle_{H^1} = \langle f, \psi \rangle, \quad \text{for all } \psi \in C_0^\infty(\Omega). \tag{6.14}$$

Note that by (6.13), $f = (-\Delta + 1)u$ in the weak sense.

By the Riesz lemma (Theorem 2.28), the condition (6.14) holds if and only if $\langle u, \cdot \rangle_{H^1}$ admits extension to $L^2(\Omega)$ as a bounded functional. We thus have an alternate specification of (6.9),

$$\mathcal{D}(-\Delta_{\mathrm{D}}) := \big\{ u \in H_0^1(\Omega) : \langle u, \cdot \rangle_{H^1} \text{ extends to } L^2(\Omega) \text{ as a bounded functional} \big\}. \tag{6.15}$$

Note that this matches the domain definition (3.29) used in the Friedrichs extension. We can thus prove a self-adjointness result by adapting the arguments from Theorem 3.29.

Theorem 6.6. *For an open set $\Omega \subset \mathbb{R}^n$, the operator $-\Delta_{\mathrm{D}}$ is self-adjoint. Moreover, it is the unique self-adjoint extension of $-\Delta$ from $C_0^\infty(\Omega)$ to a domain contained in $H_0^1(\Omega)$.*

Proof It suffices to prove the self-adjointness of $A := -\Delta + 1$, with $\mathcal{D}(A) = \mathcal{D}(-\Delta_{\mathrm{D}})$. Since A is already known to be symmetric, we need to only prove that $\mathcal{D}(A^*) \subset \mathcal{D}(A)$. The essential point will be to show that A is surjective.

Given $f \in L^2(\Omega)$, the functional $\langle f, \cdot \rangle$ is bounded on $H_0^1(\Omega)$, because

$$|\langle f, v \rangle| \le \|f\|\|v\| \le \gamma \|f\|\|v\|_{H^1}.$$

Therefore, by the Riesz lemma, there exists an element $u \in H_0^1(\Omega)$ such that

$$\langle u, v \rangle_{H^1} = \langle f, v \rangle \tag{6.16}$$

for all $v \in H_0^1(\Omega)$. By (6.15), this implies that $u \in \mathcal{D}(A)$ and

$$Au = f.$$

This proves that A is surjective. Since $\ker(A^*) = \operatorname{range}(A)^\perp$ by Lemma 3.7, this also means that A^* is injective.

Now we claim that $\mathcal{D}(A^*) \subset \mathcal{D}(A)$. Suppose that $u \in \mathcal{D}(A^*)$. By the surjectivity of A, there exists $v \in \mathcal{D}(A)$ such that

$$Av = A^* u.$$

Since $A \subset A^*$ by symmetry, this can be rewritten as

$$A^*(u - v) = 0. \tag{6.17}$$

Since A^* is injective, as noted above, this implies that $u = v$. We have thus shown that $u \in \mathcal{D}(A)$. This completes the argument that $\mathcal{D}(A^*) \subset \mathcal{D}(A)$, which proves A is self-adjoint.

To prove the uniqueness claim, suppose that B is another self-adjoint operator with $C_0^\infty(\Omega) \subset \mathcal{D}(B) \subset H_0^1(\Omega)$, such that B agrees with $-\Delta + 1$ on $C_0^\infty(\Omega)$. For $u \in \mathcal{D}(B)$ and $\psi \in C_0^\infty(\Omega)$, the self-adjointness of B implies

$$\begin{aligned} \langle Bu, \psi \rangle &= \langle u, B\psi \rangle \\ &= \langle u, (-\Delta + 1)\psi \rangle. \end{aligned}$$

By (6.10), we thus have $u \in \mathcal{D}(A) = \mathcal{D}(-\Delta_{\mathrm{D}})$ and $Bu = Au$. This shows that $B \subset A$. Since both operators are self-adjoint, we can take the adjoint of this relation to deduce $A \subset B$. Therefore $B = A$. □

Note that the uniqueness statement in Theorem 6.6 does not say that $-\Delta$ is essentially self-adjoint on $C_0^\infty(\Omega)$. There are many possible self-adjoint extensions, including the Neumann Laplacian discussed in the next section. Uniqueness holds only if we restrict our attention to domains within $H_0^1(\Omega)$.

If Ω is bounded, then the Poincaré inequality (Theorem 6.5) implies that $-\Delta_{\mathrm{D}}$ is itself bijective as a map $\mathcal{D}(-\Delta_{\mathrm{D}}) \to L^2(\Omega)$. This establishes an existence result for the Poisson problem: for any $f \in L^2(\Omega)$, there exists a unique weak

solution of $-\Delta u = f$ with $u \in \mathcal{D}(-\Delta_{\mathrm{D}})$. The inverse map $f \mapsto u$ is bounded, by Theorem 3.17. We will see in Section 6.2 that $(-\Delta_{\mathrm{D}})^{-1}$ is in fact a compact operator.

6.1.3 The Neumann Laplacian

The classical Neumann boundary condition, vanishing of the normal derivative on $\partial\Omega$, requires the definition of a normal vector to the boundary. Neumann boundary conditions have a natural interpretation, for example, in the case where $\partial\Omega$ is at least piecewise C^1. It turns out that such regularity assumptions are not required for the Neumann self-adjoint extension of $-\Delta$, just as in the Dirichlet case. However, later we will see that certain aspects of the spectral theory do require extra boundary regularity in the Neumann case.

As in Section 6.1.2, we will construct the self-adjoint extension for Neumann boundary conditions using the Friedrichs method from Section 3.4.3. The new feature here is that the boundary conditions do not appear explicitly in the specification of the quadratic form domain.

The *Neumann Laplacian* $-\Delta_{\mathrm{N}}$ is defined by the domain

$$\mathcal{D}(-\Delta_{\mathrm{N}}) := \big\{u \in H^1(\Omega) : \langle u, \cdot\rangle_{H^1} \text{ extends to } L^2(\Omega) \\ \text{as a bounded functional}\big\}. \tag{6.18}$$

This is quite similar to (6.15), except that H_0^1 is replaced by H^1. As in the Dirichlet case, the condition that $\langle u, \cdot\rangle_{H^1}$ admits a bounded L^2 extension implies that $-\Delta u \in L^2(\Omega)$. What is new here is that the boundary condition is implicit in the condition that $\langle u, \cdot\rangle_{H^1}$ is a bounded functional.

To see how this works, consider the case when $\partial\Omega$ is piecewise C^1. For $u, f \in C^\infty(\overline{\Omega})$, Green's identity gives

$$\langle \nabla u, \nabla f\rangle = \langle -\Delta u, f\rangle + \int_{\partial\Omega} f \frac{\partial \overline{u}}{\partial \nu}\, dS. \tag{6.19}$$

If we also assume that $u \in \mathcal{D}(-\Delta_{\mathrm{N}})$, then this means that

$$\langle u, v\rangle_{H^1} = \langle(-\Delta + 1)u, v\rangle$$

for all $v \in H^1(\Omega)$. This implies in particular that

$$\langle \nabla u, \nabla f\rangle = \langle -\Delta u, f\rangle$$

for all $f \in C^\infty(\overline{\Omega})$. Therefore, by (6.19),

$$\left.\frac{\partial u}{\partial \nu}\right|_{\partial\Omega} = 0$$

for $u \in \mathcal{D}(-\Delta_N) \cap C^\infty(\overline{\Omega})$.

The proof of self-adjointness in the Neumann case is essentially a repeat of the proof of Theorem 6.6. We simply replace the Hilbert space $H_0^1(\Omega)$ with $H^1(\Omega)$, to obtain the following:

Theorem 6.7. *For an open set $\Omega \subset \mathbb{R}^n$, the operator $-\Delta_N$ is self-adjoint.*

To make a uniqueness statement as in Theorem 6.6, we need for $\partial\Omega$ to be regular enough that classical Neumann boundary conditions make sense. For example, if $\partial\Omega$ is piecewise C^1 and we define the set of functions satisfying classical Neumann conditions as

$$C_N^\infty(\overline{\Omega}) := \left\{ f \in C^\infty(\overline{\Omega}) : \frac{\partial u}{\partial \nu} = 0 \right\},$$

then an argument similar to that given in Theorem 6.6 shows that $-\Delta_N$ is essentially self-adjoint on $C_N^\infty(\overline{\Omega})$.

6.2 Discreteness of Spectrum

A standard way to see that an operator T has discrete spectrum is to show that $(T - z)^{-1}$ is compact for some $z \in \rho(T)$. As we saw in Exercise 4.11, this either holds at all points in the resolvent set or none. In the former case, T is said to have *compact resolvent*, and this implies that $\sigma(T)$ is purely discrete. For a self-adjoint operator we can go even farther. By Theorem 4.21 (Hilbert–Schmidt), a self-adjoint operator with compact resolvent admits an orthonormal basis of eigenvectors.

In this section we will establish the compactness of the resolvent for the Dirichlet and Neumann Laplacians in the case where Ω is bounded. In the Neumann case the compact resolvent property requires an additional restriction on the boundary regularity. The set Ω is said to have *Lipschitz boundary* if $\partial\Omega$ can be represented locally as the graph of a Lipschitz continuous function. From the compact resolvent property we obtain the following:

Theorem 6.8. *For a bounded open set $\Omega \subset \mathbb{R}^n$, there exists an orthonormal basis $\{\psi_k\}$ for $L^2(\Omega)$, consisting of real-valued eigenfunctions of $-\Delta_D$, with real eigenvalues $\{\lambda_k\}$ accumulating at $+\infty$. The same result holds for $-\Delta_N$ if $\partial\Omega$ is Lipschitz.*

The Poincaré inequality (Theorem 6.5) implies that the eigenvalues of $-\Delta_D$ are strictly positive. If $u \in H_0^1(\Omega)$ is expanded in terms of the Dirichlet eigenfunction basis as $\sum_{k=1}^\infty c_k \psi_k$, then

$$\|\nabla u\|^2 = \sum_{k=1}^{\infty} \lambda_k |c_k|^2.$$

Assuming that the Dirichlet eigenvalues $\{\lambda_k\}$ are written in increasing order, it follows that

$$\min_{u \in H_0^1(\Omega)\setminus\{0\}} \frac{\|\nabla u\|^2}{\|u\|^2} = \lambda_1. \tag{6.20}$$

In other words, λ_1 is related to the optimal Poincaré constant γ by

$$\lambda_1 = \gamma^{-2}.$$

We will generalize (6.20) to a min–max formula for higher eigenvalues in Section 6.4.

The remainder of this section is devoted to the proof of Theorem 6.8. The main tool used to establish compactness of the resolvent is the following embedding theorem that relates the H^1 and L^2 topologies.

Theorem 6.9 (Rellich's Theorem). *For a bounded open set $\Omega \subset \mathbb{R}^n$, the inclusion $H_0^1(\Omega) \to L^2(\Omega)$ is compact. If $\partial\Omega$ is Lipschitz, then the inclusion $H^1(\Omega) \to L^2(\Omega)$ is also compact.*

Before getting into the details of the proof of Theorem 6.9, let us show that Rellich's theorem implies the compact resolvent condition.

Proof of Theorem 6.8 The proofs in Dirichlet and Neumann cases are nearly identical, so we consider only the Dirichlet case.

If $u \in \mathcal{D}(-\Delta_{\mathrm{D}})$, then by (6.21), Using (6.12), and the fact that $C_0^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, we can deduce that

$$\langle -\Delta u, v\rangle = \langle \nabla u, \nabla v\rangle = \langle u, -\Delta v\rangle \tag{6.21}$$

for $u, v \in \mathcal{D}(-\Delta_{\mathrm{D}})$. This shows that $-\Delta_{\mathrm{D}}$ is symmetric and positive.

$$\|u\|_{H^1}^2 = \langle u, (-\Delta + 1)u\rangle.$$

By Cauchy–Schwarz and the fact that $\|u\| \le \|u\|_{H^1}$, this implies that

$$\|u\|_{H^1} \le \|(-\Delta + 1)u\|.$$

This shows that $(-\Delta_{\mathrm{D}} + 1)^{-1}$ is bounded as a map $L^2(\Omega) \to H_0^1(\Omega)$. Therefore $(-\Delta_{\mathrm{D}} + 1)^{-1}$ is compact as a map $L^2(\Omega) \to L^2(\Omega)$ by Theorem 6.9. □

6.2.1 Periodic Sobolev Spaces

It is relatively easy to prove Rellich's theorem in the case of periodic functions, by making use of Fourier series. We will develop the periodic case in this section, and then later apply this to the proof of Theorem 6.9. To discuss periodic functions on $\mathbb{R}^n$, we use the notation $\mathbb{T}^n$ to denote the quotient $(\mathbb{R}/2\pi\mathbb{Z})^n$. A function on $\mathbb{T}^n$ is represented by a function on $\mathbb{R}^n$ which is 2π-periodic in each coordinate.

There are two possibilities for the definition of Sobolev spaces on $\mathbb{T}^n$. The first is to adapt the definition of weak derivatives to the periodic setting and use the analog of (2.23). Because $\mathbb{T}^n$ is compact and has no boundary, we do not need for the test functions to be restricted to compact support here. For $u \in L^1(\mathbb{T}^n)$, the weak derivative $D^\alpha u \in L^1(\mathbb{T}^n)$ is defined by the condition that

$$\int_{\mathbb{T}^n} \psi D^\alpha u \, dx = (-1)^{|\alpha|} \int_{\mathbb{T}^n} u D^\alpha \psi \, dx$$

for all $\psi \in C^\infty(\mathbb{T}^n)$. (Here α is a multi-index; see Section 2.5.1 for the introduction to this notation.) The Sobolev space $H^m(\mathbb{T}^n)$ is given by

$$H^m(\mathbb{T}^n) := \left\{ u \in L^2(\mathbb{T}^n) : \; D^\alpha u \in L^2(\mathbb{T}^n) \text{ for } |\alpha| \le m \right\}, \tag{6.22}$$

for $m \in \mathbb{N}$.

The second option is to define Sobolev spaces on $\mathbb{T}^n$ using the discrete Fourier transform. The Fourier basis for $L^2(\mathbb{T}^n)$ is defined by

$$\phi_k(x) := (2\pi)^{-n/2} e^{ik\cdot x},$$

for $k \in \mathbb{Z}^n$. The Fourier coefficient map

$$u \mapsto \hat{u}(k) := \langle \phi_k, u \rangle$$

yields an isomorphism $L^2(\mathbb{T}^n) \to \ell^2(\mathbb{Z}^n)$, as explained in Example 2.32.

It is easy to see that $D^\alpha u$ exists as a weak derivative in $L^2(\mathbb{T}^n)$ if and only if $k^\alpha \hat{u}(k) \in \ell^2(\mathbb{Z}^n)$, and that

$$D^\alpha u = \sum_{k \in \mathbb{Z}^n} (ik)^\alpha \hat{u}(k) \phi_k$$

in this case. Thus, an equivalent definition to (6.22) is

$$H^m(\mathbb{T}^n) = \left\{ u \in L^2(\mathbb{T}^n) : \sum_{k \in \mathbb{Z}^n} (1 + |k|^2)^m |\hat{u}(k)|^2 < \infty \right\},$$

which now extends the definition to all $m \ge 0$.

Since each ϕ_k is an eigenfunction of $-\Delta$ with eigenvalue $|k|^2$, it is clear that $-\Delta$ is self-adjoint on $L^2(\mathbb{T}^n)$, with the domain $H^2(\mathbb{T}^n)$.

Lemma 6.10. *For $m > 0$, the operator $(-\Delta + 1)^{-m}$ is compact on $L^2(\mathbb{T}^n)$.*

Proof For $N > 0$, define a finite rank approximation to $(-\Delta + 1)^{-m}$ by

$$T_N u = \sum_{|k| \le N} \left(|k|^2 + 1\right)^{-m} \hat{u}(k)\phi_k.$$

By the estimate

$$\left\|(-\Delta + 1)^{-m} - T_N\right\| \le \left(N^2 + 1\right)^{-m},$$

T_N converges to $(-\Delta + 1)^{-m}$ in operator norm as $N \to \infty$. It follows that $(-\Delta + 1)^{-m}$ is compact, by Theorem 3.37. □

The analog of Rellich's theorem for the periodic case is a direct consequence of Lemma 6.10.

Corollary 6.11. *The inclusion $H^m(\mathbb{T}^n) \to L^2(\mathbb{T}^n)$ is compact for $m > 0$.*

Proof Suppose that $\{w_k\}$ is a bounded sequence in $H^m(\mathbb{T}^n)$. This is equivalent to the condition that the sequence $\{(-\Delta + 1)^{m/2} w_k\}$ is bounded in $L^2(\mathbb{T}^n)$. Therefore, by the compactness of the operator $(-\Delta + 1)^{-m/2}$, there exists a subsequence of $\{w_k\}$ that converges in $L^2(\mathbb{T}^n)$. □

6.2.2 Extension Lemmas

To apply Corollary 6.11 to the proof of Theorem 6.9, we need a way to transpose Sobolev functions from Ω to $\mathbb{T}^n$. By a linear change of coordinates, it suffices to consider the case where $\overline{\Omega} \subset (0, 2\pi)^n$. The plan is to first extend H^1 functions from Ω to $(0, 2\pi)^n$ and then make them periodic.

For $H^1_0(\Omega)$ we can actually just take the naive approach of extension by zero. This works without any regularity conditions on $\partial\Omega$.

Lemma 6.12. *For open sets $\Omega \subset \widetilde{\Omega} \subset \mathbb{R}^n$, extension by zero defines an isometry $H^1_0(\Omega) \hookrightarrow H^1_0(\widetilde{\Omega})$.*

Proof For $u \in H^1_0(\Omega)$, let $\tilde{u}$ denote the extension by zero of u to a function on $\widetilde{\Omega}$. The weak first derivatives $\partial_j u \in L^2(\Omega)$ can also be extended by zero to functions $f_j \in L^2(\widetilde{\Omega})$. We must check that f_j equals the weak derivative $\partial_j \tilde{u}$.

Let $\{\psi_k\} \subset C^\infty_0(\Omega)$ be an approximating sequence such that $\psi_k \to u$ with respect to the H^1 norm. For $\phi \in C^\infty_0(\widetilde{\Omega})$, integration by parts gives

$$\int_{\widetilde{\Omega}} \phi\, \partial_j \psi_k \, d^n x = -\int_{\widetilde{\Omega}} \psi_k\, \partial_j \phi \, d^n x.$$

Taking $k \to \infty$ then shows that

$$\int_{\widetilde{\Omega}} \phi f_j \, d^n x = -\int_{\widetilde{\Omega}} \tilde{u} \, \partial_j \phi \, d^n x.$$

Hence $\partial_j \tilde{u} = f_j \in L^2(\widetilde{\Omega})$ and therefore $\tilde{u} \in H_0^1(\widetilde{\Omega})$. □

The extension problem for general H^m functions is more difficult. If Ω has a rough boundary, then it may not be possible to extend a function while maintaining the same level of regularity. To avoid technicalities in the argument, we will consider only the case when $\partial\Omega$ is smooth. Our first step is to check that the Sobolev spaces behave nicely under coordinate changes.

Lemma 6.13. *Suppose that $F : \widetilde{\Omega} \to \Omega$ is a diffeomorphism. Then the pullback $F^* : u \mapsto u \circ F$ defines a continuous bijection $H^m(\Omega) \to H^m(\widetilde{\Omega})$ for each $m \in \mathbb{N}$.*

Proof Let us represent the coordinate change by $y = F(x)$, with D_y^α and D_x^α denoting the derivatives in the respective coordinate systems. For $u \in H^m(\Omega)$, our goal is to compute $D_y^\alpha \tilde{u}$ for $|\alpha| \le m$, where $\tilde{u} := u \circ F$. For $\tilde{\psi} \in C_0^\infty(\widetilde{\Omega})$, a change of coordinates by $G := F^{-1}$ yields

$$(-1)^{|\alpha|} \int_{\widetilde{\Omega}} \tilde{u} \, D_y^\alpha \tilde{\psi} \, dy = (-1)^{|\alpha|} \int_{\Omega} u \, G^*(D_y^\alpha \tilde{\psi}) \, |\det J_G| \, dx,$$

where J_G is the Jacobian matrix of G. By the chain rule,

$$G^*(D_y^\alpha \tilde{\psi}) = \sum_{|\beta| \le |\alpha|} c_\beta D_x^\beta \psi,$$

for some coefficients $c_\beta \in C^\infty(\Omega)$. Since $u \in H^m(\Omega)$, this implies that

$$(-1)^{|\alpha|} \int_{\widetilde{\Omega}} \tilde{u} \, D_y^\alpha \tilde{\psi} \, dy = \int_{\Omega} v^{(\alpha)} \psi \, dx,$$

where $v^{(\alpha)} \in L^2(\Omega)$ is given by

$$v^{(\alpha)} := \sum_{|\beta| \le |\alpha|} D_x^\beta \big(c_\beta |\det J_G| \, u\big).$$

It follows that the weak derivative $D_y^\alpha \tilde{u}$ exists and is equal to $F^* v^{(\alpha)} \in L^2(\widetilde{\Omega})$. Furthermore, since F^* is bounded as a map $L^2(\Omega) \to L^2(\widetilde{\Omega})$,

$$\|D_y^\alpha \tilde{u}\| \le C \|v^{(\alpha)}\| \le C \|u\|_{H^m}.$$

□

Using Lemma 6.13, we can simplify the extension problem by first flattening the boundary with a change of coordinates. This strategy yields the following:

Lemma 6.14 (Sobolev Extension). *Suppose that $\Omega \subset \mathbb{R}^n$ has smooth boundary. A function in $u \in H^m(\Omega)$ admits an extension to $\tilde{u} \in H^m(\mathbb{R}^n)$, such that*

$$\|\tilde{u}\|_{H^m(\mathbb{R}^n)} \leq C_m \|u\|_{H^m(\Omega)}, \tag{6.23}$$

with C_m independent of u.

Proof The first step is to reduce this to a local argument. By assumption, each boundary point has a neighborhood in which $\partial\Omega$ is the graph of a smooth function. Since $\partial\Omega$ is compact, we can cover it with a finite number of such neighborhoods. A smooth partition of unity can then be used to restrict the estimate to one such neighborhood. (For partition of unity in $\mathbb{R}^n$ see, e.g., Rudin [77, Thm. 10.8].)

It thus suffices to consider the case where $\operatorname{supp} u \in U \subset \mathbb{R}^n$, where $\partial\Omega \cap U$ is the graph of a smooth function. After an interchange of coordinates, if necessary, we can assume that

$$U \cap \Omega = \{x \in U : x_n > \vartheta(x_1, \ldots, x_{n-1})\},$$

for some smooth function ϑ.

We can now apply a simple coordinate change on U to straighten the boundary. For $x \in U$ define the map $y = F(x)$ by,

$$y_j = \begin{cases} x_j, & j = 1, \ldots, n-1, \\ x_n - \vartheta(x_1, \ldots, x_{n-1}), & j = n, \end{cases} \tag{6.24}$$

so that $F(U \cap \Omega)$ is the set $\{y_n \geq 0\}$, as illustrated in Figure 6.4. The inverse map is given by $x_n = y_n + \vartheta(y_1, \ldots, y_{n-1})$, so it is clear that F is a diffeomorphism.

By Lemma 6.13, it suffices to prove the result in the new coordinates. We can thus specialize to the case where u is a compactly supported function in $H^m(\mathbb{R}^n_+)$, where $\mathbb{R}^n_+ = \{x_n > 0\}$. We need to extend u across $x_n = 0$ to a function $\tilde{u} \in H^m(\mathbb{R}^n)$.

For this construction, suppose first that $u \in C_0^\infty(\overline{\mathbb{R}^n_+})$. Such a function admits a smooth extension across the boundary, by definition. The point here is to control this extension so that (6.23) holds. The strategy is to use rescaled reflections of u.

For a set of coefficients $c_1, \ldots, c_{m+1} \in \mathbb{C}$, consider the ansatz,

$$\tilde{u}(x) := \begin{cases} u(x), & x_n \geq 0, \\ \displaystyle\sum_{j=1}^{m+1} c_j u\left(x_1, \ldots, x_{n-1}, -\frac{x_n}{j}\right), & x_n < 0. \end{cases} \tag{6.25}$$

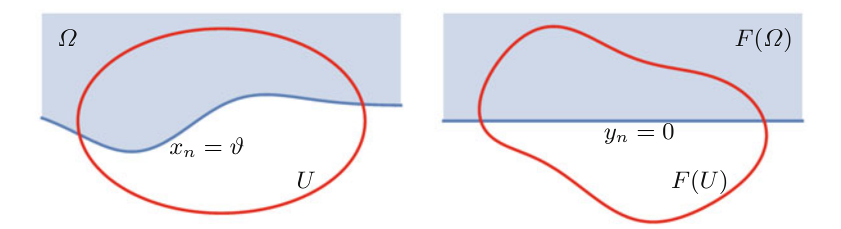


Fig. 6.4 Using a change of coordinates to flatten the boundary

The function $\tilde{u}$ is continuous provided $\sum c_j = 1$, and this condition also implies that partial derivatives involving only the variables $x_1, \ldots, x_{n-1}$ are continuous across the boundary.

To match up the x_n derivatives, we compute

$$\frac{\partial^l \tilde{u}}{\partial x_n^l}\Big|_{x_n \to 0^-} = \sum_{j=1}^{m+1} (-j)^{-l} c_j \frac{\partial^l u}{\partial x_n^l}\Big|_{x_n=0}.$$

After comparing this to (6.25), we see that $\tilde{u} \in C^m(\mathbb{R}^n)$ under the condition

$$\sum_{j=1}^{m+1} (-j)^{1-i} c_j = 1, \tag{6.26}$$

for $i = 1, \ldots, m+1$. This linear system has coefficient matrix

$$M_{ij} := (-j)^{-i+1},$$

and to show that (6.26) has a solution we simply need to know that $\det M \neq 0$. The matrix M is of Vandermonde type, and a simple induction argument gives the determinant formula,

$$\det M = \sum_{1 \le i < j \le m+1} \left(\frac{1}{i} - \frac{1}{j} \right),$$

which is nonzero. Therefore, there exists a unique choice of $c_1, \ldots, c_{m+1}$ satisfying (6.26). For example, for $m = 1$ the solution is $c_1 = -3$ and $c_2 = 4$.

Using the set of coefficients that solves (6.26) in (6.25) gives $\tilde{u} \in C_0^m(\mathbb{R}^n)$. Moreover, since the c_j depend only on m, it is clear that (6.23) holds for all $u \in C_0^\infty(\overline{\mathbb{R}_+^n})$.

Now suppose $u \in H^m(\mathbb{R}^n_+)$. To make use of the extension constructed above, we need to approximate u by smooth functions, a process called *mollification*. Suppose $\psi \in C_0^\infty(\mathbb{R}^n)$ has support in $\{|x| < 1\}$ and satisfies

$$\int_{\mathbb{R}^n} \psi \, d^n x = 1. \tag{6.27}$$

For $x \in \overline{\mathbb{R}^n_+}$ and $\varepsilon > 0$, we define a shifted convolution,

$$u_\varepsilon(x) := \int_{|y|<1} \psi(y) u(x + \varepsilon e_n - \varepsilon y)\, d^n y,$$

where e_n is the unit vector in the x_n direction. (The shift in the x_n variable keeps the argument of u within $\mathbb{R}^n_+$.) By the dominated convergence theorem, $u_\varepsilon \to u$ in $L^2(\mathbb{R}^n_+)$ as $\varepsilon \to 0$.

Moreover, a simple change of variables gives

$$u_\varepsilon(x) = \int_{|y|<1} \psi((x - y)/\varepsilon + e_n) u(y)\, d^n y.$$

We can then differentiate under the integral to see that $u_\varepsilon \in C^\infty(\overline{\mathbb{R}^n_+})$, with

$$\begin{aligned} D^\alpha u_\varepsilon(x) &= \int_{|y|<1} D_x^\alpha \Big[\psi\Big(\varepsilon^{-1}(x - y) + e_n\Big)\Big] u(y)\, d^n y \\ &= (-1)^{|\alpha|} \int_{|y|<1} D_y^\alpha \Big[\psi\Big(\varepsilon^{-1}(x - y) + e_n\Big)\Big] u(y)\, d^n y. \end{aligned}$$

Since $D^\alpha u$ exists as a weak derivative for $|\alpha| \le m$, this implies

$$\begin{aligned} D^\alpha u_\varepsilon(x) &= \int_{|y|<1} \psi\Big(\varepsilon^{-1}(x - y) + e_n\Big) D^\alpha u(y)\, d^n y \\ &= \int_{|y|<1} \psi(y) D^\alpha u(x + \varepsilon e_n - \varepsilon y)\, d^n y. \end{aligned}$$

This shows that $D^\alpha u_\varepsilon \to D^\alpha u$ in $L^2(\mathbb{R}^n)$ for $|\alpha| \le m$. Hence $u_\varepsilon \to u$ in $H^m(\mathbb{R}^n_+)$.

Now let $\tilde{u}_\varepsilon$ be the extension of u_ε to $\mathbb{R}^n$ as defined above. Since we have already shown that (6.23) applies to $u_\varepsilon \in C^\infty(\overline{\mathbb{R}^n_+})$, the convergence $u_\varepsilon \to u$ implies that the sequence $\{\tilde{u}_\varepsilon\}$ is Cauchy in $H^m(\mathbb{R}^n)$. Therefore, the limit

$$\tilde{u} := \lim_{\varepsilon \to 0} \tilde{u}_\varepsilon$$

exists in $H^m(\mathbb{R}^n)$.

As noted above, this local argument can be applied to $u \in H^m(\Omega)$ using a partition of unity. □

It is easy to see that the extension result of Lemma 6.14 actually requires only that $\partial\Omega$ is C^m. In fact, with a more careful argument, this boundary requirement can be relaxed to the assumption that $\partial\Omega$ is of the Hölder class $C^{m-1,1}$; see Gilbarg and Trudinger [36, Thm. 7.25]. In particular, for $m = 1$ this means that the boundary is required to be Lipschitz continuous.

To conclude this section, we use Lemma 6.14, in conjunction with Corollary 6.11, to complete the proof of Rellich's theorem.

Proof of Theorem 6.9 First consider the embedding of $H^1_0(\Omega)$ into $L^2(\Omega)$. By translating and rescaling if needed, we can assume that $\overline{\Omega} \subset (0, 2\pi)^n$. This allows us to extend a compactly supported function on $(0, 2\pi)^n$ to periodic function on $\mathbb{T}^n$. Using this periodic extension along with the extension property of Lemma 6.12 gives an isometry

$$H^1_0(\Omega) \to H^1(\mathbb{T}^n). \tag{6.28}$$

Suppose $\{v_k\} \subset H^1_0(\Omega)$ is a bounded sequence, and let $\{\tilde{v}_k\}$ denote the corresponding periodic extensions in $H^1(\mathbb{T}^n)$. By Corollary 6.11, there exists a subsequence of $\{\tilde{v}_k\}$ converging in $L^2(\mathbb{T}^n)$. The corresponding subsequence of $\{v_k\}$ then converges in $L^2(\Omega)$.

The argument is almost the same for $H^1(\Omega)$, except that the extension result requires an extra hypothesis. If $\partial\Omega$ is smooth, this extension is covered by Lemma 6.14. For the more general condition that $\partial\Omega$ is Lipschitz, we cite the stronger extension result mentioned above [36, Thm. 7.25]. □

6.3 Regularity of Eigenfunctions

The eigenfunctions provided by Theorem 6.8 are required to solve $-\Delta\psi = \lambda\psi$ only in the weak sense. In this section, we will show that the weak eigenvalue equation requires ψ to be smooth, so that eigenfunctions are in fact classical solutions.

This argument, a special case of elliptic regularity from PDE theory, is relatively easy if we restrict our attention to the interior of Ω. Suppose ψ is an eigenfunction for either the Dirichlet or Neumann Laplacian on a bounded open set $\Omega \subset \mathbb{R}^n$. For $\chi \in C^\infty_0(\Omega)$, we can extend $\chi\psi$ by zero to a function in $H^1(\mathbb{R}^n)$. By the eigenvalue equation for u,

$$-\Delta(\chi\psi) = \lambda\chi\psi - [\Delta, \chi]\psi. \tag{6.29}$$

The commutator $[\Delta, \chi]$ is a differential operator of order one, so the right-hand side of (6.29) lies in $L^2(\mathbb{R}^n)$. The Fourier transform of the left-hand side is $|\xi|^2\widehat{\chi\psi}$. By (2.26), this shows that $\chi\psi \in H^2(\mathbb{R}^n)$. This puts the right-hand side of (6.29) in $H^1(\mathbb{R}^n)$. Iterating this argument shows that $\chi\psi \in H^m(\mathbb{R}^n)$ for all m, which implies $\chi\psi \in C^\infty(\mathbb{R}^n)$ by Theorem 2.26. Since χ was arbitrary, this shows that

$$\psi \in C^\infty(\Omega). \tag{6.30}$$

It follows from Lemma 2.22 that ψ satisfies the eigenvalue equation in the classical sense.

In fact, we can go further than smoothness. A function $\Omega \to \mathbb{R}$ is *real analytic* if it is locally representable by power series. That is, the Taylor series at each point converges to the function in a neighborhood. Harmonic functions (solutions of $\Delta u = 0$) are real analytic, and this property extends to eigenfunctions of the Laplacian.

Theorem 6.15. *For an open set $\Omega \subset \mathbb{R}^n$, suppose that $\psi \in H^1(\Omega)$ satisfies $-\Delta\psi = \lambda\psi$ in the weak sense, for $\lambda \in \mathbb{R}$. Then ψ is a real analytic function on Ω.*

Before proving this result, let us note that both (6.30) and Theorem 6.15 are only concerned with regularity in the interior. It is possible to extend regularity results up to the boundary, but this requires a corresponding degree of regularity of $\partial\Omega$. For example, if $\partial\Omega$ is smooth, then the eigenfunction satisfies $\psi \in C^\infty(\overline{\Omega})$. We will defer this boundary regularity discussion to the more general context of Riemannian manifolds; see Section 9.4.2 and Appendix A.4.

The existence of a local power series representation implies that if a real analytic function vanishes to infinite order at a point, then it vanishes identically on the entire connected component of the domain. The proof is the same as for holomorphic functions in complex analysis. Thus Theorem 6.15 has the following:

Corollary 6.16 (Unique Continuation). *Assume that $\Omega \subset \mathbb{R}^n$ is a connected open set. If $\psi \in \Omega$ satisfies $-\Delta\psi = \lambda\psi$ and vanishes to infinite order at a point in Ω, then $\psi \equiv 0$.*

The unique continuation property holds more generally for solutions of elliptic partial differential equations, even in the nonanalytic case, by a result of Aronszajn [2].

Our strategy for the proof of Theorem 6.15 is based on the construction of a real analytic integral kernel for $(-\Delta_{\mathbb{R}^n} - \lambda)^{-1}$ for $\lambda > 0$. This is closely related to the resolvent kernel construction in Section 4.1.3, where we produced a Green's function $G(\kappa; r)$ for $-\Delta + \kappa^2$ acting on $\mathbb{R}^n$, with $\operatorname{Re}\kappa > 0$. Here we seek a kernel function $G(ik; r)$ for $k > 0$, which will satisfy

$$\begin{aligned}\phi(x) &= (-\Delta - k^2)\int_{\mathbb{R}^n} G(ik; |x-y|)\phi(y)\, d^n y \\ &= \int_{\mathbb{R}^n} G(ik; |x-y|)(-\Delta - k^2)\phi(y)\, d^n y,\end{aligned} \tag{6.31}$$

for $\phi \in C_0^\infty(\mathbb{R}^n)$. Note that the fact that $\lambda = k^2$ lies in the spectrum does not rule out a solution of (6.31). It just means that $G(ik; r)$ cannot be the kernel of a bounded operator on L^2.

Lemma 6.17. *For $\lambda > 0$ and $x \in \mathbb{R}^n$, the function*

$$G(ik; r) := -\frac{\pi}{2}(2\pi)^{-\frac{n}{2}}\left(\frac{r}{k}\right)^{1-\frac{n}{2}} Y_{\frac{n}{2}-1}(kr), \tag{6.32}$$

where $Y_\nu(z)$ denotes the Bessel function, satisfies (6.31).

Proof We start by noting that, because the action of $(-\Delta - k^2)$ is local, the first line of (6.31) implies that

$$(-\Delta - k^2)G(ik; |x|) = 0$$

for $x \neq 0$. As a radial equation, this translates to

$$\frac{\partial^2 G}{\partial r^2} + \frac{n-1}{r}\frac{\partial G}{\partial r} + k^2 G = 0.$$

for $r > 0$. By setting $z = kr$ and $G(ik; r) = z^{-\nu} f(z)$ with $\nu := n/2 - 1$, we can reduce this to the Bessel equation

$$z^2 f'' + z f' + (z^2 - \nu^2) f = 0. \tag{6.33}$$

The solutions are linear combinations of the Bessel functions $J_\nu(z)$ and $Y_\nu(z)$. The function $|x|^{-\nu} J_\nu(k|x|)$ is smooth at $x = 0$. We can see from the calculation in Section 4.1.3 that the Green's function should have a singularity at the origin. Therefore, we make the ansatz

$$G(ik; r) := a z^{-\nu} Y_\nu(z), \tag{6.34}$$

for some constant a.

The two integral formulas in (6.31) are equivalent for $\phi \in C_0^\infty(\mathbb{R}^n)$, by integration by parts, so we can focus on the second. By a change of variables it suffices to prove this at the origin, i.e.,

$$\phi(0) = \int_{\mathbb{R}^n} G(ik; |y|)(-\Delta - k^2)\phi(y) d^n y. \tag{6.35}$$

Since (6.34) is locally integrable, we can rewrite the integral as a limit

$$\int_{\mathbb{R}^n} G(ik;|y|)(-\Delta-k^2)\phi(y)d^ny = \lim_{\varepsilon\to 0}\int_{\{r\geq\varepsilon\}} G(ik;|y|)(-\Delta-k^2)\phi(y)\,d^ny.$$

Note that the integral avoids the singularity at the origin, so we can integrate by parts, using Green's identity, to move the operator from ϕ onto G. Since $(-\Delta - k^2)G = 0$ away from the origin, the integral reduces to a boundary term,

$$\int_{\mathbb{R}^n} G(r;|y|)(-\Delta-k^2)\phi(y)\,d^ny = \lim_{\varepsilon\to 0}\int_{\{r=\varepsilon\}}\left(G\frac{\partial\phi}{\partial r} - \phi\frac{\partial G}{\partial r}\right)dS. \tag{6.36}$$

As $z \to 0$, the Bessel Y-function satisfies the asymptotic [64, §10.7],

$$Y_\nu(z) \sim \begin{cases} -\dfrac{2^\nu}{\pi}\Gamma(\nu)z^{-\nu}, & \operatorname{Re}\nu > 0 \text{ or } \nu\in -\mathbb{N}+\frac{1}{2},\\ \dfrac{2}{\pi}\log z, & \nu = 0. \end{cases} \tag{6.37}$$

This shows that $G(\lambda;r) = O(r^{2-n})$ for $n\neq 2$ and $O(\log r)$ for $n = 2$, implying

$$\lim_{\varepsilon\to 0}\int_{\{r=\varepsilon\}} G\frac{\partial\phi}{\partial r}\,dS = 0, \tag{6.38}$$

since the surface measure is $O(r^{n-1})$. Furthermore, by the derivative formula,

$$\frac{\partial}{\partial z}\big[z^{-\nu}Y_\nu(z)\big] = -z^{-\nu}Y_{\nu+1}(z),$$

we deduce from (6.37) that

$$\frac{\partial G}{\partial r} \sim -\frac{1}{2}\pi^{-\frac{n}{2}}\Gamma(\tfrac{n}{2})r^{1-n}.$$

Since

$$\operatorname{vol}(\mathbb{S}^{n-1}) = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})},$$

this implies that

$$\lim_{\varepsilon\to 0}\int_{\{r=\varepsilon\}}\phi\frac{\partial G}{\partial r}dS = -\phi(0). \tag{6.39}$$

Applying (6.38) and (6.39) to the right-hand side of (6.36) proves (6.35). □

From (6.37), we can see that the behavior of $G(ik; r)$ as $r \to 0$ is identical to that of $G(\kappa; r)$. The reason that $G(ik; r)$ does not correspond to a bounded operator is the lack of sufficient decay at infinity. For example, in $\mathbb{R}^3$ we have

$$G(ik; r) = \frac{\cos(kr)}{4\pi r}.$$

More generally. one can check that $G(ik; r)$ grows like $\log r$ in dimension two and decays like r^{1-n} in dimension $n \geq 3$.

In terms of the regularity of eigenfunctions, the key point in (6.35) is that $G(ik; |x|)$ is an analytic function of x for $x \neq 0$. This follows from the fact that $Y_\nu(\cdot)$ is holomorphic on the domain $\mathbb{C} \backslash (-\infty, 0]$.

Proof of Theorem 6.15 Suppose $\psi \in H^1(\Omega)$ satisfies $-\Delta\psi = k^2\psi$, in the weak sense. Let V be an open set such that $\overline{V} \subset \Omega$, and choose a cutoff function $\chi \in C_0^\infty(\Omega)$ so that $\chi = 1$ on V. We have already observed that ψ is smooth in (6.30), so the function,

$$\begin{aligned} f &:= (-\Delta - k^2)(\chi\psi) \\ &= -[\Delta, \chi]\psi, \end{aligned}$$

is smooth and supported within $\operatorname{supp}(\nabla\chi)$. In particular, f vanishes on V. By Lemma 6.17, for $x \in V$,

$$\psi(x) = \int_{\mathbb{R}^n} G(ik; |x - y|) f(y)\, d^n y. \tag{6.40}$$

Since $G(ik; |x - y|)$ is analytic as a function of $x \in V$ for $y \in \operatorname{supp}(f)$, this implies that ψ is analytic on V. □

6.4 Eigenvalue Computations

For a bounded open set $\Omega \subset \mathbb{R}^n$, let $\{\psi_k\}_{k=1}^\infty$ denote the set of eigenfunctions corresponding to the Dirichlet eigenvalues (6.4). For $u \in H_0^1(\Omega)$, the fact that the eigenfunctions constitute an orthonormal basis for $L^2(\Omega)$ implies that

$$\|\nabla u\|^2 = \sum_{k=1}^\infty \lambda_k\, |\langle e_k, u\rangle|^2.$$

This shows that

$$\lambda_1 = \min_{u \in H_0^1(\Omega)\backslash\{0\}} \frac{\|\nabla u\|^2}{\|u\|^2}. \tag{6.41}$$

The ratio on the right is called the *Rayleigh quotient* of u. By comparing (6.41) to the Poincaré inequality Theorem 6.5, we see that λ_1 is related to the Poincaré constant γ by

$$\lambda_1 = \gamma^{-2}.$$

The characterization of λ_1 in (6.41) differs from the corresponding expression for λ_1 from Theorem 5.15 in that the minimum is taken over the quadratic form domain $H_0^1(\Omega)$, rather than the operator domain. It is not difficult to extend the quadratic form version to a full statement of the min–max principle.

Theorem 6.18 (Min–Max Principle for Dirichlet Eigenvalues). *Let $\{\lambda_k\}$ be the set of Dirichlet eigenvalues of a bounded open set $\Omega \subset \mathbb{R}^n$, written in increasing order and repeated according to multiplicity. Define Λ_k as the set of subspaces of $H_0^1(\Omega)$ of dimension k. Then*

$$\lambda_k = \min_{W \in \Lambda_k} \left\{ \max_{u \in W \setminus \{0\}} \frac{\|\nabla u\|^2}{\|u\|^2} \right\} \tag{6.42}$$

for each $k \in \mathbb{N}$.

Proof Because the quadratic form domain $H_0^1(\Omega)$ contains the operator domain $\mathcal{D}(-\Delta)$, Theorem 5.15 implies that

$$\lambda_k \geq \min_{W \in \Lambda_k} \left\{ \max_{u \in W \setminus \{0\}} \frac{\|\nabla u\|^2}{\|u\|^2} \right\}. \tag{6.43}$$

To prove the opposite inequality, consider a general subspace $W \in \Lambda_k$. Since $\dim W = k$, there exists a nonzero vector

$$w \in W \cap [\psi_1, \ldots, \psi_{k-1}]^{\perp},$$

where $\{\psi_j\}$ denotes the eigenvalue basis. This implies the eigenfunction expansion,

$$w = \sum_{j=k}^{\infty} \langle w, \psi_j \rangle \psi_j.$$

Therefore,

$$\|\nabla w\|^2 = \sum_{j=k}^{\infty} \lambda_j |\langle w, \psi_j \rangle|^2,$$

which shows that

$$\|\nabla w\|^2 \geq \lambda_k \|w\|^2.$$

Since W was arbitrary, it follows that

$$\lambda_k \leq \min_{W \in \Lambda_k} \left\{ \max_{u \in W \setminus \{0\}} \frac{\|\nabla u\|^2}{\|u\|^2} \right\}. \tag{6.44}$$

In combination with (6.43), this completes the proof. □

With a similar argument, we can develop a max–min formula, analogous to (5.31):

$$\lambda_k = \max_{v_1,\ldots,v_{k-1} \in L^2(\Omega)} \left\{ \min_{u \in H_0^1(\Omega) \cap \{v_1,\ldots,v_{k-1}\}^\perp \setminus \{0\}} \frac{\|\nabla u\|^2}{\|u\|^2} \right\}. \tag{6.45}$$

We will discuss the corresponding formulas for Neumann eigenvalues in Section 6.4.3.

6.4.1 Finite Element Method

Theorem 6.18 is often used to approximate eigenvalues by choosing a finite-dimensional subspace $\mathcal{A} \subset H_0^1(\Omega)$ and computing the min–max values for the restriction to $\mathcal{A}$,

$$\beta_k := \min_{W \in \Lambda_k(\mathcal{A})} \left\{ \max_{u \in W \setminus \{0\}} \frac{\|\nabla u\|^2}{\|u\|^2} \right\}, \tag{6.46}$$

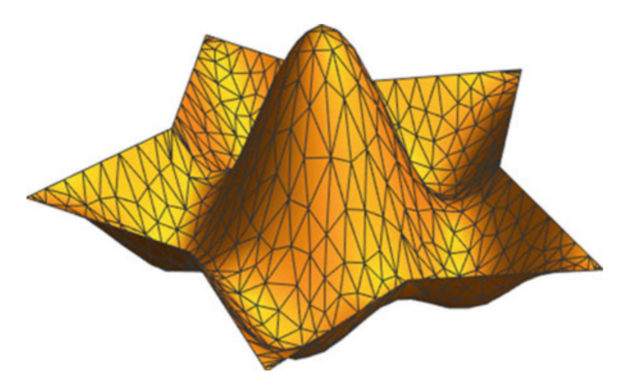
for $k = 1, \ldots, m$, where $m = \dim \mathcal{A}$ and $\Lambda_k(\mathcal{A})$ denotes the subspaces of $\mathcal{A}$ of dimension k. This computational technique is known as the *Rayleigh–Ritz method*. Clearly $\lambda_k \leq \beta_k$ for each k, since $\Lambda_k(\mathcal{A}) \subset \Lambda_k(\mathcal{H})$.

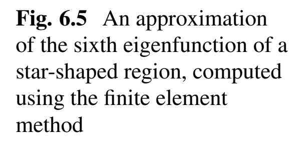
To compute β_k, we choose a basis $\{w_k\}$ for $\mathcal{A}$, not necessarily orthonormal. Then define the matrices

$$E_{ij} := \langle \nabla w_i, \nabla w_j \rangle, \qquad F_{ij} = \langle w_i, w_j \rangle,$$

for $i, j = 1, \ldots, m$. A simple linear algebra argument shows that the β_k are the eigenvalues of the matrix $F^{-1}E$. The idea behind the Rayleigh–Ritz method is that β_k should be a good approximation to λ_k, at least if k is small relative to m. Moreover, if $(c_1, \ldots, c_m)$ denotes the eigenvector of $F^{-1}E$ corresponding to β_k, then $\sum c_j w_j$ can be used as an approximation to the true eigenfunction ϕ_k.

Example 6.19. Consider the Dirichlet Laplacian on the interval $[0, 1]$, for which the spectrum is $\{(k\pi)^2\}_{k=1}^\infty$. The Rayleigh quotient of the polynomial $p_1(x) = x(1-x)$ is 10, a somewhat crude approximation to π^2. Suppose instead we apply the Rayleigh–Ritz method with $\mathcal{A}$ given by the span of polynomials $p_j(x) = x^j(1-x)$




Fig. 6.5 An approximation of the sixth eigenfunction of a star-shaped region, computed using the finite element method

for $j = 1, \dots, 5$. Then the first Rayleigh–Ritz value, $\beta_1 \doteq 9.8696$, approximates π^2 to within 10^{-7}. The second value, $\beta_2 \doteq 39.50$, is still a reasonable approximation to $4\pi^2 \doteq 39.48$. ♢

For the Rayleigh–Ritz method to be effective, we need an efficient way to choose a subspace $\mathcal{A}$ that approximates $H^1_0(\Omega)$ reasonably well. A standard approach is the *finite element method*, which involves subdividing Ω into a polygonal mesh and then constructing $\mathcal{A}$ using continuous, piecewise linear functions ("elements") which take the value 1 at a single vertex of the mesh and vanish at all others. These sample elements lie in $H^1_0(\Omega)$, and the corresponding matrices E and F can be computed quite efficiently, since most of the entries are zero. Figure 6.5 shows an approximate eigenfunction computed using a mesh with 883 triangles.

6.4.2 Domain Monotonicity

The min–max principle proves to be very useful for eigenvalue comparisons. For example, we have the following monotonicity property for nested domains.

Theorem 6.20. *For a pair of bounded open subsets of $\mathbb{R}^n$ satisfying $\Omega \subset \widetilde{\Omega}$, the Dirichlet eigenvalues satisfy*

$$\lambda_k(\Omega) \geq \lambda_k(\widetilde{\Omega})$$

for each k.

Proof For any subspace $\tilde{W} \subset H^1_0(\widetilde{\Omega})$ of dimension k we have

$$\lambda_k(\widetilde{\Omega}) \leq \max_{u \in \tilde{W} \setminus \{0\}} \frac{\|\nabla u\|^2}{\|u\|^2}, \tag{6.47}$$

by (6.42). Let $\{\psi_k\}$ denote the eigenfunctions of $-\Delta_{\mathrm{D}}$ on Ω. By Lemma 6.12, these functions interpreted as elements of $H^1_0(\widetilde{\Omega})$ by extension by zero. Setting $\tilde{W}$ equal to $\operatorname{span}\{\psi_1, \dots, \psi_k\}$ in (6.47) then gives the result. □

The counting function for Dirichlet eigenvalues is defined as

$$N_\Omega(t) := \#\big\{\lambda_k \in \sigma : \lambda_k \le t\big\}. \tag{6.48}$$

Theorem 6.20 immediately yields a rough estimate of $N_\Omega(t)$, by comparison to the counting function for a domain whose eigenvalues are known, such as the rectangular domain from Example 6.1. If $\mathcal{R}_1 \subset \Omega \subset \mathcal{R}_2$, then

$$N_{\mathcal{R}_2}(t) \le N_\Omega(t) \le N_{\mathcal{R}_1}(t). \tag{6.49}$$

To put this to use, we need to compute the asymptotic behavior of $N_{\mathcal{R}}(t)$ as $t \to \infty$. Let ω_n be the volume of the unit ball in $\mathbb{R}^n$,

$$\omega_n := \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}. \tag{6.50}$$

Lemma 6.21. *The Dirichlet eigenvalue counting function for a rectangular region $\mathcal{R}$ satisfies the asymptotic*

$$N_{\mathcal{R}}(t) \sim (2\pi)^{-n}\omega_n \operatorname{vol}(\mathcal{R}) t^{\frac{n}{2}}, \tag{6.51}$$

as $t \to \infty$. This implies the eigenvalue asymptotic,

$$\lambda_k \sim (2\pi)^2 \left(\frac{k}{\omega_n \operatorname{vol}(\mathcal{R})}\right)^{2/n}, \tag{6.52}$$

as $k \to \infty$.

Proof If the side lengths of $\mathcal{R}$ are labeled $\ell_1, \dots \ell_n$, then by Example 6.1,

$$N_{\mathcal{R}}(t) = \#\left\{\nu \in \mathbb{N}^n : \sum_{j=1}^n \frac{\pi^2 \nu_j^2}{\ell_j^2} \le t\right\}.$$

To estimate this counting function, consider the elliptical sector

$$E(r) := \left\{x \in (\mathbb{R}_+)^n : \sum_{j=1}^n \frac{\pi^2 x_j^2}{\ell_j^2} \le r^2\right\}.$$

Each lattice point counted in $N_{\mathcal{R}}(t)$ corresponds to a unit cube contained in $E(\sqrt{t})$. The union of these cubes is illustrated as the shaded region in Figure 6.6. This inclusion gives the upper bound

$$N_{\mathcal{R}}(t) \le \operatorname{vol} E(\sqrt{t}).$$

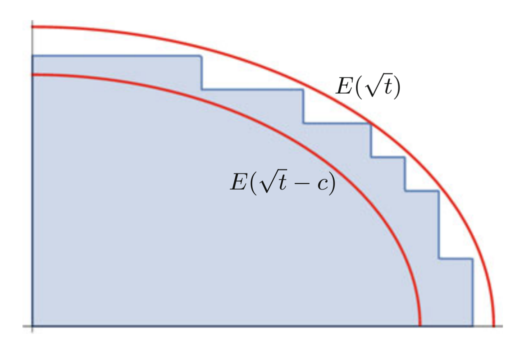


Fig. 6.6 Estimating the counting function $N_{\mathcal{R}}(t)$

On the other hand, $E(\sqrt{t} - c)$ is contained within the union of cubes for $c = \pi\sqrt{n}/\min(\ell_j)$, so that

$$N_{\mathcal{R}}(t) \geq \text{vol}\, E(\sqrt{t} - c).$$

The volume of the elliptical sector is

$$\text{vol}\, E(r) = \omega_n r^n \prod_{j=1}^{n} \frac{\ell_j}{2\pi}.$$

Noting that $\prod \ell_j = \text{vol}(\mathcal{R})$, we deduce that

$$N_{\mathcal{R}}(t) = (2\pi)^{-n}\omega_n\, \text{vol}(\mathcal{R}) t^{\frac{n}{2}} + O(t^{(n-1)/2}), \tag{6.53}$$

as $t \to \infty$.

The corresponding eigenvalue asymptotic (6.52) follows immediately from the observation that $\lambda_k \leq t < \lambda_{k+1}$ for $k = N_{\mathcal{R}}(t)$. □

Using (6.49) and Lemma 6.21, we can now derive a growth estimate for $N_\Omega(t)$. The symbol $\asymp$ means that the ratio of the two sides is bounded above and below by strictly positive constants.

Corollary 6.22. *For the Dirichlet Laplacian on a bounded open set $\Omega \subset \mathbb{R}^n$, the eigenvalue counting function satisfies*

$$N_\Omega(t) \asymp t^{n/2}.$$

Equivalently,

$$\lambda_k \asymp k^{2/n}.$$

We will see in Section 6.5 that the crude estimate of Corollary 6.22 can be refined into an asymptotic of the form (6.51) for any bounded open set. This formula is called *Weyl's law*.

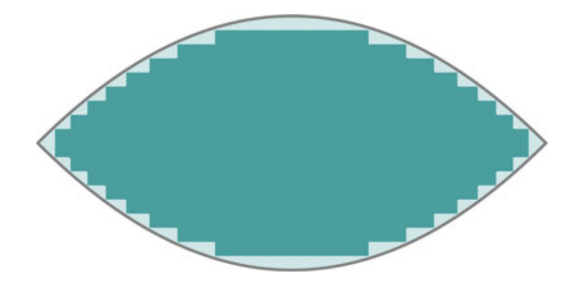


Fig. 6.7 Approximating a domain from within by a finite union of rectangles

One way to obtain this refinement is by noting that the comparison principle of Theorem 6.20 does not require the domains to be connected. Suppose we take a finite disjoint union of rectangles $\mathcal{Q}_j$ contained in Ω, as illustrated in Figure 6.7. Then

$$\lambda_k(\Omega) \leq \lambda_k(\cup_{j=1}^m \mathcal{Q}_j). \tag{6.54}$$

The upper bound in (6.49) can thus be replaced by

$$N_\Omega(t) \leq \sum_{j=1}^m N_{\mathcal{Q}_j}(t).$$

By taking families of disjoint rectangles that approximate Ω arbitrarily well in terms of Lebesgue measure, we can derive from this method a sharp estimate of $\limsup_{t\to\infty} N_\Omega(t)$.

6.4.3 Neumann Eigenvalues

There is an important relationship between Dirichlet and Neumann eigenvalues that proves useful in the analysis of both cases. It follows immediately from the Neumann case of the min–max principle.

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with boundary satisfying the regularity assumption of Theorem 6.8, i.e., $\partial\Omega$ is at least Lipschitz continuous. To avoid confusion with the Dirichlet case, we write the Neumann eigenvalues as

$$0 = \mu_1 \leq \mu_2 \leq \mu_3 \leq \dots,$$

where eigenvalues are repeated according to multiplicity. (The zero eigenvalue could be repeated only if Ω has more than one connected component.)

The same arguments used in the proof of Theorem 6.18, with $H_0^1(\Omega)$ replaced by $H^1(\Omega)$, yield the following:

Theorem 6.23 (Min–Max Principle for Neumann Eigenvalues). *For a $\Omega \subset \mathbb{R}^n$ a bounded open set with boundary at least Lipschitz continuous, let Λ_k denote the set of subspaces of $H^1(\Omega)$ of dimension k. The Neumann eigenvalues of $-\Delta$ are given by*

$$\mu_k = \min_{W \in \Lambda_k} \left\{ \max_{u \in W \setminus \{0\}} \frac{\|\nabla u\|^2}{\|u\|^2} \right\} \tag{6.55}$$

for each $k \in \mathbb{N}$.

Note that the Neumann version of the min–max formula includes a larger class of subspaces than the Dirichlet. It follows immediately that the Neumann eigenvalues will be smaller.

Corollary 6.24 (Dirichlet–Neumann Comparison). *For a $\Omega \subset \mathbb{R}^n$ as above, the Dirichlet and Neumann eigenvalues of $-\Delta$ are related by*

$$\mu_k \leq \lambda_k,$$

for all $k \in \mathbb{N}$.

Despite this direct comparison between Dirichlet and Neumann eigenvalues, the domain monotonicity property established for the Dirichlet case in Theorem 6.20 does not hold in the Neumann case. The argument from the Dirichlet case does not apply here because elements of $H^1(\Omega)$ do not generally admit extensions that preserve the Rayleigh quotient. We can easily find counterexamples that show the failure of monotonicity.

Example 6.25. Let Ω be the unit square in $\mathbb{R}^2$. For $a \in (0, 1]$, define a rectangle $\mathcal{R}$ inside Ω with vertices $(a, 0)$, $(0, a)$, $(1-a, 1)$, and $(1, 1-a)$, as shown in Figure 6.8. Suppose the Neumann eigenvalues of each domain are written as

$$0 = \mu_1 \leq \mu_2 \leq \mu_2 \leq \dots.$$

For $a < \frac{1}{2}$, we have

$$\mu_2(\Omega) = \pi^2, \quad \mu_2(\mathcal{R}) = \frac{\pi^2}{2(1-a)^2}.$$

Thus $\mu_2(\mathcal{R})$ could be either larger or smaller than $\mu_2(\Omega)$, depending on whether a is greater or less than $1 - \frac{1}{\sqrt{2}}$. ♢

There is a very restrictive form of the domain monotonicity result that holds for Neumann eigenvalues. To see this, consider the max–min formula,

$$\mu_k = \max_{v_1,\dots,v_{k-1} \in L^2(\Omega)} \left\{ \min_{u \in H^1(\Omega) \cap \{v_1,\dots,v_{k-1}\}^\perp \setminus \{0\}} \frac{\|\nabla u\|^2}{\|u\|^2} \right\}. \tag{6.56}$$

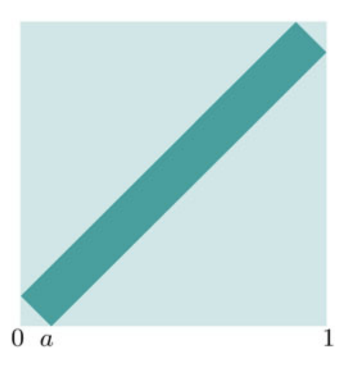


Fig. 6.8 Counterexample for Neumann domain multiplicity

If $\widetilde{\Omega}\backslash\Omega$ has measure zero, then $H^1(\widetilde{\Omega})$ can be regarded as subspace of $H^1(\Omega)$ by restriction. Moreover, $L^2(\Omega) = L^2(\widetilde{\Omega})$ under this assumption. Since the restriction map does not change the Rayleigh quotient, we obtain from (6.56) the following:

Lemma 6.26 (Restricted Neumann Domain Monotonocity). *Suppose that $\Omega \subset \widetilde{\Omega}$ are open subsets of $\mathbb{R}^n$ with Lipchitz continuous boundary, such that $\widetilde{\Omega}\backslash\Omega$ has measure zero. Then*

$$\mu_k(\Omega) \leq \mu_k(\widetilde{\Omega})$$

for each k.

A typical case for Lemma 6.26 is the subdivision of a domain by adding internal boundaries. For example, a domain with rectangular sides can be cut into a finite union of rectangles. As noted in Section 6.4.2, such a subdivision causes the Dirichlet eigenvalues to increase, but by Lemma 6.26 the Neumann eigenvalues would decrease.

This idea can be used to provide a corresponding lower bound to (6.54). Suppose $\{\mathcal{R}_i\}$ is a finite collection of disjoint rectangles, such that Ω is contained in the union of the closures of the $\mathcal{R}_i$. By the combination of Theorem 6.20, Corollary 6.24, and Lemma 6.26, we have

$$\lambda_k(\Omega) \geq \mu_k(\cup\mathcal{R}_i). \tag{6.57}$$

Estimating eigenvalues by comparison to rectangles, using a combination of (6.54) and (6.57), is referred to as *Dirichlet–Neumann bracketing*.

6.5 Asymptotics of Dirichlet Eigenvalues

In this section we will develop a proof of Weyl's law, the asymptotic formula for the eigenvalue counting function $N_\Omega(t)$ mentioned in Section 6.4.2. This asymptotic had been conjectured independently by Lorentz and Sommerfeld in 1910.

Hermann Weyl proved it in 1911, for two-dimensional domains with sufficiently regular boundary, using the Dirichlet–Neumann bracketing method described in the preceding section.

Theorem 6.27 (Weyl's Law). *For the Dirichlet Laplacian on a bounded open set $\Omega \subset \mathbb{R}^n$,*

$$N_\Omega(t) \sim (2\pi)^{-n}\omega_n \operatorname{vol}(\Omega) t^{\frac{n}{2}}$$

as $t \to \infty$, where ω_n is the volume of the unit ball in $\mathbb{R}^n$, (6.50). Equivalently, if the eigenvalues are arranged in increasing order

$$\lambda_k \sim (2\pi)^2 \left(\frac{k}{\omega_n \operatorname{vol}(\Omega)} \right)^{2/n}$$

as $k \to \infty$.

Example 6.28. In the early nineteenth century, Gabriel Lamé [56] computed the spectrum of an equilateral triangle using trigonometric polynomials. For an equilateral triangle $\mathcal{T}$ with side length ℓ, a complete set of Dirichlet eigenvalues (with multiplicities) is given by

$$\lambda_{k,m} = \frac{16\pi^2}{9\ell^2}(k^2 + km + m^2),$$

for $k, m \in \mathbb{N}$. A sample eigenfunction is shown in Figure 6.9.

We can interpret $N_\mathcal{T}(t)$ as the count of integer lattice points within the elliptical sector

$$E(t) := \left\{ x \geq 0,\ y \geq 0,\ \frac{16\pi^2}{9\ell^2}(x^2 + xy + y^2) \leq t \right\}.$$

As in Lemma 6.21, this implies that

$$N_\mathcal{T}(t) \sim \operatorname{area} E(t).$$

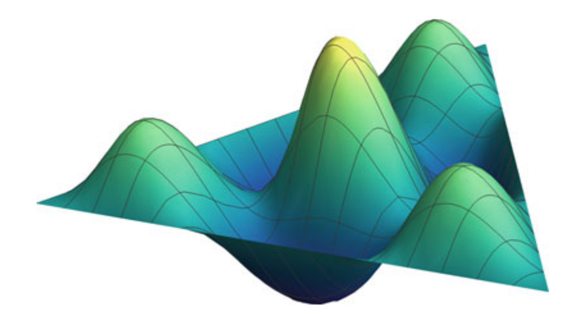


Fig. 6.9 A Lamé eigenfunction on the equilateral triangle

In polar coordinates, $x^2 + xy + y^2 = r^2(1 + \frac{1}{2}\sin 2\theta)$, which allows us to compute

$$\begin{aligned}\operatorname{area} E(t) &= \frac{9\ell^2}{16\pi^2 t}\int_0^{\pi/2} \frac{d\theta}{2 + \sin 2\theta} \\ &= \frac{\sqrt{3}\ell^2}{4}\frac{t}{4\pi}.\end{aligned}$$

Since $\mathcal{T}$ has area $\sqrt{3}\ell^2/4$, this agrees with the asymptotic stated in Theorem 6.27. $\diamondsuit$

One obvious consequence of Weyl's law is the fact that the volume of Ω is fixed by its spectrum. This naturally invites the question of what other geometric properties of Ω are determined by the spectrum. For example, Åke Pleijel proved in 1954 [67] that the length of $\partial\Omega$ is determined by the spectrum if Ω is a two-dimensional domain with smooth boundary.

The question of whether Ω is completely determined by its spectrum was memorably formulated in a famous 1966 article by Mark Kac [48], titled "Can you hear the shape of a drum?" (Kac credits Lipman Bers as the source of this eloquent phrasing.) The question was answered in the negative by Carolyn Gordon, David Webb, and Scott Wolpert [38] in 1992. An example of a pair of "isospectral" domains is shown in Figure 6.10. Many other polygonal counterexamples have been found. However, for domains with smooth boundary, Kac's question remains open.



Fig. 6.10 Bounded open sets in $\mathbb{R}^2$ with the same Dirichlet eigenvalues

6.5.1 Strategy for the Proof

Although Weyl considered only the two-dimensional case of Theorem 6.27, his Dirichlet–Neumann bracketing argument generalizes easily to higher dimensions. However, this approach does impose a regularity assumption on $\partial\Omega$, arising from the approximation of Ω by finite collections of rectangles. We can approximate $\operatorname{vol}(\Omega)$ from the inside by a finite union of open rectangles, to arbitrary precision with respect to Lebesgue measure. Similarly, we can approximate $\operatorname{vol}(\overline{\Omega})$ to arbitrary precision with a covering by finitely many closed rectangles. The problem

is that these inner and outer volumes may not be equal. The bracketing argument is restricted to sets for which $\mathrm{vol}(\partial\Omega) = 0$, which are called *Jordan measurable*.

A different approach to the Weyl asymptotic was introduced by Torsten Carleman [18] in 1934, based on an asymptotic analysis of the resolvent kernel of the Dirichlet Laplacian. Carleman's method paved the way for more modern approaches to Weyl asymptotics, involving the heat and wave operators. It also yields local information about the distribution of eigenfunctions; see Exercise 6.5.

Lars Gårding [34] later applied Carleman's approach to the case of arbitrary uniformly elliptic operators, even possibly non-self-adjoint. Our presentation in this section essentially follows Gårding's, although we restrict our attention to the Laplacian to simplify the exposition.

The argument is based on analysis of the asymptotic behavior of the resolvent of $-h^2\Delta + 1$ as $h \to 0$. The use of the letter h here reflects the fact that the Schrödinger operator in quantum mechanics has $-\hbar^2\Delta$ as its leading term, where $\hbar$ is Planck's constant. The "correspondence principle" of quantum mechanics says that classical physics should be recovered by rescaling units so that $\hbar \to 0$. Because of this connection, studying the limiting behavior of an operator such as $-h^2\Delta + 1$ as $h \to 0$ is called *semiclassical* analysis. (See, e.g., Zworski [97].)

The full proof of the Weyl law is rather involved and will be spread out into a few subsections. The major steps are as follows:

Section 6.5.2 We work out the asymptotics of the kernel of $(-h^2\Delta + 1)^{-m}$ as $h \to 0$, for $m \in \mathbb{N}$.

Section 6.5.3 For m sufficiently large, the operator $(-h^2\Delta + 1)^{-m}$ is trace-class. The pointwise asymptotics of its kernel are used to derive a limiting formula for the trace of $(-h^2\Delta + 1)^{-m}$ as $h \to 0$.

Section 6.5.4 The resolvent trace is written as a sum over eigenvalues, and the Weyl formula is deduced from the asymptotic behavior of this sum as $h \to 0$, using a Tauberian theorem.

6.5.2 Asymptotics of the Resolvent Kernel

On a bounded domain $\Omega \subset \mathbb{R}^n$, we define the differential operator

$$P(h) := -h^2\Delta + 1$$

for $h > 0$. For the proof of the Weyl formula, we will need to distinguish between this differential operator and its self-adjoint extensions on various domains. Let $A(h)$ denote the realization of $P(h)$ as a self-adjoint operator on $L^2(\Omega)$ with Dirichlet boundary conditions, defined as in Section 6.1, with $\mathcal{D}(A(h)) = \mathcal{D}(-\Delta) \subset H^1_0(\Omega)$. Let $A_0(h)$ denote the self-adjoint extension of $P(h)$ to $L^2(\mathbb{R}^n)$, with $\mathcal{D}(A_0(h)) = H^2(\mathbb{R}^n)$.

Note that the inverses $A(h)^{-1}$ and $A_0(h)^{-1}$ are proportional to the respective resolvents, $(-\Delta_D + h^{-2})^{-1}$ and $-(\Delta_{\mathbb{R}^n} + h^{-2})^{-1}$. These inverses therefore exist for all $h > 0$ as bounded operators on $L^2(\Omega)$ and $L^2(\mathbb{R}^n)$, respectively.

The main point of Carleman's strategy is to analyze the semiclassical limit of a trace over the eigenvalues of $A(h)$. Although $A(h)^{-1}$ is not trace-class, we can form a related trace-class operator by taking the power $A(h)^{-m}$ for m sufficiently large. Our goal is to express the leading term in the trace of $A(h)^{-m}$, as $h \to 0$, in terms of the kernel of $A_0(h)^{-m}$.

As a first step, let us work out the explicit formula for the kernel of $A_0(h)^{-m}$. The case $m = 1$ was already covered in Section 4.1.3, where the result was expressed in terms of the modified Bessel function,

$$K_\nu(z) := \frac{1}{2}\left(\frac{z}{2}\right)^\nu \int_0^\infty t^{\nu-1} e^{-t-z^2/4t}\, dt, \tag{6.58}$$

for $\nu \in \mathbb{R}$ and $z > 0$.

Lemma 6.29. *For $\phi \in C_0^\infty(\mathbb{R}^n)$ and $m > 0$,*

$$A_0(h)^{-m}\phi(x) = \int_{\mathbb{R}^n} \Phi_m(h; |x-y|)\phi(y)\, d^n y,$$

where

$$\Phi_m(h; r) = \frac{2^{1-m}(2\pi)^{-n/2}}{\Gamma(m)} h^{-m-\frac{n}{2}} r^{m-\frac{n}{2}} K_{m-\frac{n}{2}}(r/h),$$

and K_ν is given by (6.58).

Proof Since the action of $A_0(h)$ is conjugate to multiplication by $h^2|\xi|^2 + 1$ under the Fourier transform,

$$A_0(h)^{-m}\phi(x) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} \frac{\hat{\phi}(\xi)}{(h^2|\xi|^2+1)^m}\, d^n\xi. \tag{6.59}$$

To compute the integral over ξ, we substitute

$$(h^2|\xi|^2+1)^{-m} = \frac{1}{\Gamma(m)} \int_0^\infty t^{m-1} e^{-t(h^2|\xi|^2+1)}\, dt$$

into (6.59), and then use

$$\int_{\mathbb{R}^n} e^{ix\cdot\xi - th^2|\xi|^2}\, d\xi = \left(\frac{\pi}{th^2}\right)^{\frac{n}{2}} e^{-|x|^2/4th^2}.$$

Applying these computations to (6.59) yields

$$\Phi_m(h;r) = \frac{(4\pi)^{-\frac{n}{2}}}{\Gamma(m)} h^{-n} \int_0^\infty t^{m-\frac{n}{2}-1} e^{-t-r^2/4th^2}\, dt.$$

The claimed formula then follows by setting $z = r/h$ and $\nu = m - n/2$ in the definition (6.58). □

The next step is to relate the kernel of $A(h)^{-m}$ to $\Phi_m(h;r)$. The crucial observation for this purpose is that $A(h)$ and $A_0(h)$ are extensions of the same differential operator $P(h)$. In the following argument, we will consider a more general extension $\tilde{P}(h)$, such that $A(h) \subset \tilde{P}(h)$. We will eventually need to apply this result not just to $A(h)$, but also to the Dirichlet extension of $P(h)$ on a larger bounded open set containing Ω.

Theorem 6.30. *Let $\Omega \subset \mathbb{R}^n$, and suppose that $\tilde{P}(h)$ is a self-adjoint extension of $P(h)$ with a domain that includes $C_0^\infty(\Omega)$. For a fixed $m \in \mathbb{N}$, there exists a kernel function $\varrho(h;\cdot,\cdot) \subset C^\infty(\Omega \times \Omega)$ such that*

$$\left\langle \phi_1, \tilde{P}(h)^{-m}\phi_2 \right\rangle = \left\langle \phi_1, A_0(h)^{-m}\phi_2 \right\rangle + \int_\Omega \int_\Omega \varrho(h;x,y)\overline{\phi_1(y)}\phi_2(y')\, d^n y\, d^n y'$$

for $\phi_1, \phi_2 \in C_0^\infty(\Omega)$, with

$$\varrho(h;x,y) = O(h^\infty),$$

as $h \to 0$, uniformly on compact subsets of $\Omega \times \Omega$.

Proof Our goal is to find an integral formula for the sesquilinear form on $C_0^\infty(\Omega)$ defined by

$$\eta[\phi_1,\phi_2] := \left\langle \phi_1, \left(\tilde{P}(h)^{-m} - A_0(h)^{-m}\right)\phi_2 \right\rangle.$$

Note that, because both $\tilde{P}(h)$ and $A_0(h)$ reduce to $P(h)$ on $C_0^\infty(\Omega)$, we have

$$\eta\left[P(h)^m\phi_1, \phi_2\right] = \eta\left[\phi_1, P(h)^m\phi_2\right] = 0, \tag{6.60}$$

for all $\phi_1, \phi_2 \in C_0^\infty(\Omega)$.

Fix an open set V such that $\overline{V} \subset \Omega$, and assume now that $\phi_1, \phi_2 \in C_0^\infty(V)$. Choose $\chi \in C_0^\infty(\Omega)$ such that $\chi = 1$ on some open neighborhood of $\overline{V}$. We can deduce from (6.59) that the functions $A_0(h)^{-m}\phi_j$ are smooth, by the arguments used to prove Sobolev embedding (Theorem 2.26). Thus, using (6.60) and the fact that $P(h)^m A_0^{-m}\phi_j = \phi_j$, we have the identity,

$$\eta[\phi_1,\phi_2] = \eta\Big[P(h)^m(1-\chi)A_0(h)^{-m}\phi_1, P(h)^m(1-\chi)A_0(h)^{-m}\phi_2\Big]. \tag{6.61}$$

By the formula from Lemma 6.29,

$$P(h)^m(1-\chi)A_0(h)^{-m}\phi_j(x) = \int_V b(x,y)\phi_j(y)\, d^n y, \tag{6.62}$$

where

$$b(x,y) := (-h^2\Delta_x + 1)^m\Big[(1-\chi(x))\Phi_m(h;|x-y|)\Big], \tag{6.63}$$

for $x \in \mathbb{R}^n$, $y \in V$. The function $b(\cdot,y)$ is smooth for each $y \in V$, because $\Phi_m(h,r)$ is smooth for $r > 0$ and $1-\chi$ vanishes in some neighborhood of $\overline{V}$. Furthermore, for $y \in V$ we have $\operatorname{supp} b(\cdot,y) \subset \operatorname{supp}(\nabla\chi)$, because

$$(-h^2\Delta_x + 1)^m\Phi_m(h;|x-y|) = 0 \tag{6.64}$$

for $x \neq y$.

Therefore, we can combine (6.62) with (6.61) and apply Fubini's theorem to deduce that

$$\eta[\phi_1,\phi_2] = \int_V\int_V \varrho(h;y,y')\overline{\phi_1(y)}\phi_2(y')\, d^n y\, d^n y', \tag{6.65}$$

where

$$\varrho(h;y,y') := \eta[b(\cdot,y), b(\cdot,y')]. \tag{6.66}$$

By the properties of $b(\cdot,\cdot)$ noted above, the kernel $\varrho(h;\cdot,\cdot)$ is smooth on $V \times V$.

Since $\|\tilde{P}(h)^{-1}\| \le 1$ and $\|A_0(h)^{-1}\| = 1$, we obtain from (6.66) the estimate

$$|\varrho(h;y,y')| \le 2\|b(\cdot,y)\|\,\|b(\cdot,y')\|. \tag{6.67}$$

The Bessel function $K_\nu(z)$ and its derivatives are $O(e^{-z})$ as $z \to +\infty$. This means that $\Phi_m(h;|x-y|)$ and its derivatives can be estimated by $O(e^{-c/h})$ for $c > 0$, as long as $|x-y|$ is bounded away from zero. By the definition (6.63), $\operatorname{supp} b(\cdot,\cdot) \subset \operatorname{supp}(\nabla\chi) \times V$. Since $\operatorname{dist}(\operatorname{supp}(\nabla\chi), V) > 0$, we can deduce from (6.63) that

$$\sup_{x,y} |b(x,y)| = O(e^{-c/h})$$

for some $c > 0$. Hence (6.67) gives the estimate

$$\sup_{y,y'\in V} |\varrho(h;y,y')| = O(e^{-c/h}). \tag{6.68}$$

The constants in this estimate depend on χ and V.

Suppose we apply the same construction on a larger subset V_1, to produce a kernel function ϱ_1. The fact that (6.65) holds for both ϱ and ϱ_1, provided $\phi_1, \phi_2 \subset C_0^\infty(V)$, shows that ϱ_1 agrees with ϱ on $V \times V$. Thus, there exists a unique extension of ϱ to a smooth function on $\Omega \times \Omega$. From (6.68) we can derive the claimed $O(h^\infty)$ bound on ϱ in any compact subset of $\Omega \times \Omega$. □

The lack of uniformity in Theorem 6.30 is a technical problem that we will address in our study of trace asymptotics in Section 6.5.3. At this point we can only use the result to derive a local version of the asymptotic. Let $\{\psi_k\}$ denote the eigenfunctions of $-\Delta_{\mathrm{D}}$ on Ω, corresponding to the eigenvalues $\{\lambda_k\}$. Each ψ_k is also an eigenfunction of $A(h)^{-m}$, with

$$A(h)^{-m}\psi_k = (h^2\lambda_k + 1)^{-m}\psi_k.$$

Setting $\tilde{P}(h) = A(h)$ in Theorem 6.30 yields the following:

Corollary 6.31. *For $m > n/2$,*

$$\sum_{k=1}^{\infty} \frac{\psi_k(x)\overline{\psi_k(y)}}{(h^2\lambda_k + 1)^m} = \begin{cases} (4\pi)^{-\frac{n}{2}} \dfrac{\Gamma(m - \frac{n}{2})}{\Gamma(m)} h^{-n} + O(h^{-n+1}), & x = y, \\ O(h^\infty), & x \neq y, \end{cases}$$

uniformly on compact subsets of $\Omega \times \Omega$ as $h \to 0$.

Proof By the estimate $\lambda_k \asymp k^{2/n}$ from Corollary 6.22, the sum

$$K_m(h; x, y) := \sum_{k=1}^{\infty} \frac{\psi_k(x)\overline{\psi_k(y)}}{(h^2\lambda_k + 1)^m} \tag{6.69}$$

converges in $L^2(\Omega \times \Omega)$ for $m > n/4$. This defines an integral kernel for $A(h)^{-m}$, which is therefore a Hilbert–Schmidt operator.

By Theorem 6.30, the integral kernel could also be written as

$$K_m(h; x, y) = \Phi_m(h; |x - y|) + \varrho(h; x, y),$$

where Φ_m is the kernel function computed in Lemma 6.29. For the Bessel function $K_\nu(z)$, we can deduce from (6.58) the asymptotic as $z \to 0$,

$$K_\nu(z) \sim 2^{\nu-1}\Gamma(\nu)z^{-\nu}, \tag{6.70}$$

for $\nu > 0$. Therefore, $\Phi_m(h; r)$ is continuous at $r = 0$ for $m > n/2$. In this case, $K_m(h; \cdot, \cdot)$ is continuous on $\Omega \times \Omega$. Mercer's theorem (Theorem 4.23) then shows that the expansion (6.69) converges uniformly on compact sets.

It now follows from Theorem 6.30 that

$$K_m(h; x, y) = \Phi_m(h; |x - y|) + O(h^\infty),$$

for $m > n/2$, uniformly on compact sets. If $|x - y|$ is bounded away from zero, then the $O(h^\infty)$ estimate follows from the exponential decay of $K_\nu(z)$ as $z \to +\infty$. For the leading term on the diagonal, we can deduce from (6.70) that

$$\Phi_m(h; 0) = (4\pi)^{-\frac{n}{2}} \frac{\Gamma(m - \frac{n}{2})}{\Gamma(m)} h^{-n}$$

for $m > n/2$. □

6.5.3 *Trace Asymptotics*

To complete the proof of Theorem 6.27, we need to establish the analog of Corollary 6.31 for

$$\operatorname{tr}\big[A(h)^{-m}\big] = \sum_{k=1}^{\infty} \frac{1}{(h^2\lambda_k + 1)^m},$$

with $m > n/2$. By Theorem 4.24, the trace can be computed as the integral of the restriction of the kernel to the diagonal,

$$\operatorname{tr}\big[A(h)^{-m}\big] = \int_\Omega K_m(h; x, x) d^n x, \tag{6.71}$$

where $K_m(h; \cdot, \cdot)$ is defined by (6.69). We cannot simply integrate the asymptotic from Corollary 6.31, however, because that result is not uniform at the boundary.

Theorem 6.32. *Let $\{\lambda_k\}$ be the Dirichlet eigenvalues of a bounded open set $\Omega \subset \mathbb{R}^n$. For $m > n/2$,*

$$\sum_{k=1}^{\infty} \frac{1}{(h^2\lambda_k + 1)^m} \sim (4\pi)^{-\frac{n}{2}} \frac{\Gamma(m - \frac{n}{2})}{\Gamma(m)} \operatorname{vol}(\Omega) h^{-n}$$

as $h \to 0$.

Proof For convenience, set

$$c_{m,n} := (4\pi)^{-\frac{n}{2}} \frac{\Gamma(m - \frac{n}{2})}{\Gamma(m)},$$

By Fatou's lemma, multiplying (6.71) by h^n and using the limit from Corollary 6.31 gives

$$\liminf_{h \to 0} \left[h^n \sum_{k=1}^{\infty} \frac{1}{(h^2\lambda_k + 1)^m} \right] \geq c_{m,n} \operatorname{vol}(\Omega). \tag{6.72}$$

In other words, we have established the lower bound half of the claimed asymptotic.

For the bound from above, we must work around the lack of uniformity in the estimate of $\varrho(h;\cdot,\cdot)$ in Theorem 6.30. For this purpose, let us consider a slightly larger bounded open set $\widetilde{\Omega} \subset \mathbb{R}^n$, such that $\widetilde{\Omega} \supset \overline{\Omega}$. Let $B(h)$ denote the Dirichlet extension of $P(h)$ on $\widetilde{\Omega}$.

We claim that

$$\left\langle f, A(h)^{-1} f\right\rangle \leq \left\langle f, B(h)^{-1} f\right\rangle \tag{6.73}$$

for $f \in L^2(\Omega)$, where on the right-hand side f is interpreted as an element of $L^2(\widetilde{\Omega})$ after extending by zero. To show this, let $\beta[\cdot,\cdot]$ denote the inner product on $H_0^1(\Omega)$ associated with $B(h)$,

$$\beta[u,v] := \langle u, B(h)v\rangle.$$

For $f \in L^2(\Omega)$,

$$\begin{aligned}\left\langle f, A(h)^{-1} f\right\rangle &= \left\langle B(h)B(h)^{-1} f, A(h)^{-1} f\right\rangle \\ &= \beta\left[B(h)^{-1} f, A(h)^{-1} f\right].\end{aligned}$$

Thus, by Cauchy–Schwarz,

$$\begin{aligned}\left\langle f, A(h)^{-1} f\right\rangle^2 &\leq \left\|B(h)^{-1} f\right\|_\beta^2 \left\|A(h)^{-1} f\right\|_\beta^2 \\ &= \left\langle f, B(h)^{-1} f\right\rangle \left\langle A(h)^{-1} f, B(h)A(h)^{-1} f\right\rangle.\end{aligned} \tag{6.74}$$

Because $A(h) \subset B(h)$, we have

$$B(h)A(h)^{-1} f = f$$

for $f \in L^2(\Omega)$. Thus (6.73) follows from (6.74).

Since the estimate (6.73) is restricted to functions on the smaller domain, in order to make use of it we need to introduce a cutoff version of $B(h)^{-1}$. Let P_Ω be the restriction map $L^2(\widetilde{\Omega}) \to L^2(\Omega)$, and define

$$Q(h) := P_\Omega B(h)^{-1},$$

as an operator on $L^2(\Omega)$. Since $B(h)^{-1}$ is positive and compact, $Q(h)$ is also positive and compact. The eigenvalues of $Q(h)$ are bounded above by those of $B(h)^{-1}$, by the max–min principle (Theorem 4.22). Hence $Q(h)^m$ is trace-class for $m > n/2$.

From (6.73), we have

$$\langle f, A(h)^{-1} f\rangle \leq \langle f, Q(h) f\rangle$$

for $f \in L^2(\Omega)$. Thus, the max–min principle also implies that the eigenvalue $(h^2\lambda_k + 1)^{-1}$ of $A(h)^{-1}$ is bounded above by the kth eigenvalue of $Q(h)$. For $m > n/2$, it follows that

$$\sum_{k=1}^{\infty} \frac{1}{(h^2\lambda_k + 1)^m} \le \operatorname{tr}\big[Q(h)^m\big]. \tag{6.75}$$

Now the goal is to analyze $\operatorname{tr}[Q(h)^m]$ in the limit $h \to 0$. Taking $\tilde{P}(h) = B(h)$ in Theorem 6.30 shows that $Q(h)^m$ is represented by an integral $q_m(h; x, y)$ satisfying the pointwise asymptotic

$$q_m(h; x, x) = c_{m,n} h^{-n} + O(h^\infty), \tag{6.76}$$

for $m > n/2$, where the $O(h^\infty)$ remainder term is smooth as a function of x. By Theorem 4.24,

$$\operatorname{tr}\big[Q(h)^m\big] = \int_\Omega q_m(h; x, x)\, dx. \tag{6.77}$$

To complete the proof, the idea is to use dominated convergence to take the limit (6.76) inside the integral in (6.77). This requires a uniform bound on $q_m(h; x, x)$.

To obtain such a bound, we will exploit the fact that the integral kernel of $Q(h)$ is the restriction to $\Omega \times \Omega$ of the kernel of $B(h)^{-1}$. Thus, applying Theorem 6.30 to $B(h)^{-1}$ on $\widetilde{\Omega}$ yields the asymptotic

$$q_1(h; x, y) = \Phi_1(h; |x-y|) + \varrho_1(h; x, y), \tag{6.78}$$

for $x \neq y$, where $\varrho_1(h; \cdot, \cdot)$ is smooth and uniformly $O(h^\infty)$ on $\Omega \times \Omega$. Theorem 6.30 applies directly to q_m only in the case $m = 1$, because P_Ω does not commute with $B(h)^{-1}$. However, we can exploit (6.78) by writing q_m as an m-fold convolution of q_1.

Using the fact that $K_{-\nu}(z) = K_\nu(z)$, we have

$$\Phi_1(h; r) = (2\pi)^{-n/2} h^{-n} (r/h)^{1-n/2} K_{n/2-1}(r/h).$$

By the asymptotic (6.70), and the fact that $K_0(t) \sim -\log t$ as $t \to 0^+$, we can thus estimate, for some $\delta > 0$ (needed only if $n = 2$),

$$|\Phi_1(h; ht)| \le C h^{-n} t^{2-n-\delta} e^{-t}$$

for all $t > 0$. If we set

$$w(z) := |z|^{2-n+\delta} e^{-|z|}, \tag{6.79}$$

then from (6.78) we can estimate

$$|q_1(h; x, y)| \le C h^{-n} w((x-y)/h), \tag{6.80}$$

where C depends on n and δ, but not on x, y or h.

Writing q_m as the m-fold convolution $q_1 * \cdots * q_1$ gives

$$q_m(h; x, x) = \int_{\Omega\times\cdots\times\Omega} q_1(h; x, y_1)\ldots q_1(h; y_{m-1}, x) \prod_{j=1}^{m-1} d^n y_j.$$

If we then set $y_j = x + hz_j$ and use the estimate (6.80), this yields

$$|q_m(h; x, x)| \leq Ch^{-n} \int_{\mathbb{R}^{n(m-1)}} w(z_1)w(z_1 - z_2)\ldots w(z_{m-1}) \prod_{j=1}^{m-1} d^n z_j. \tag{6.81}$$

The remaining integral can be estimated with a repeated application of Young's inequality (Theorem A.15),

$$\int_{\mathbb{R}^{n(m-1)}} w(z_1)w(z_1 - z_2)\ldots w(z_{m-1}) \prod_{j=1}^{m-1} d^n z_j \leq C\|w\|^m_{L^{m/(m-1)}}.$$

From (6.79), we can see that $w \in L^p(\mathbb{R}^n)$ for $p < n/(n - 2 + \delta)$. Thus, $w \in L^{m/(m-1)}(\mathbb{R}^n)$ for $m > n/2$ and the integral in (6.81) is therefore finite. This yields a uniform bound

$$|q_m(h; x, x)| \leq Ch^{-n},$$

for $x \in \Omega$.

Returning to (6.77), we can multiply by h^n and apply the dominated convergence and the pointwise limit (6.76) to deduce that

$$\lim_{h\to 0}\Big(h^n \operatorname{tr}\big[Q(h)^m\big]\Big) = c_{m,n} \operatorname{vol}(\Omega).$$

Thus, by (6.75),

$$\limsup_{h\to 0}\left[h^n \sum_{k=1}^{\infty} \frac{1}{(h^2\lambda_k + 1)^m}\right] \leq c_{m,n} \operatorname{vol}(\Omega), \tag{6.82}$$

which is the upper bound half of the asymptotic. In combination, (6.72) and (6.82) give

$$\lim_{h\to 0}\left[h^n \sum_{k=1}^{\infty} \frac{1}{(h^2\lambda_k + 1)^m}\right] = c_{m,n} \operatorname{vol}(\Omega),$$

which completes the proof. □

6.5.4 The Tauberian Argument

The final step in establishing the Weyl formula is to extract asymptotics of the sequence $\{\lambda_k\}$ from the limit in Theorem 6.32, using a *Tauberian theorem*. The name comes from a result proved by Alfred Tauber in 1897 that says if $\lim_{r\to 1^-} \sum_{n=1}^{\infty} a_n r^n = A$ and $a_n = o(1/n)$, then $\sum_{n=1}^{\infty} a_n = A$. This basic result admits a wide variety of powerful generalizations. The relevant version for our purposes is the following:

Theorem 6.33 (Karamata's Theorem). *Let μ be a measure on $[0,\infty)$, such that e^{-tx} is integrable with respect to $d\mu(x)$ for each $t > 0$. Suppose that for $\alpha > 0$,*

$$\int_0^\infty e^{-tx}\, d\mu(x) \sim At^{-\alpha}, \tag{6.83}$$

as $t \to 0^+$ (or $t \to \infty$). Then

$$\mu[0,s] \sim \frac{A}{\Gamma(\alpha+1)} s^\alpha,$$

as $s \to \infty$ (or $s \to 0^+$, respectively).

Proof The proofs are quite similar, so we consider only the case $t \to 0^+$. Define a family of measures ν_t by

$$d\nu_t(x) = t^\alpha e^{-x}\, d\mu(x/t)$$

For $c \geq 0$,

$$\int_0^\infty e^{-cx}\, d\nu_t(x) = t^\alpha \int_0^\infty e^{-(c+1)tx}\, d\mu(x).$$

Thus, the hypothesis (6.83) gives

$$\lim_{t\to 0^+} \int_0^\infty e^{-cx}\, d\nu_t(x) = A(c+1)^{-\alpha}. \tag{6.84}$$

We can express the right-hand side of (6.84) in terms of a limiting measure

$$d\nu_0(x) := \frac{A}{\Gamma(\alpha)} x^{\alpha-1} e^{-x}\, dx,$$

which satisfies

$$\int_0^\infty e^{-cx}\, d\nu_0(x) = A(c+1)^{-\alpha}.$$

By the asymptotic (6.84), we have

$$\lim_{t\to 0^+}\int_0^\infty e^{-cx}\,d\nu_t(x)=\int_0^\infty e^{-cx}\,d\nu_0(x), \tag{6.85}$$

for all $c\geq 0$.

To extend (6.85) to a broader class of functions, we introduce the algebra

$$\mathcal{A}:=\left\{\sum_{j=1}^m a_j e^{-c_j x}\,:\,a_j\in\mathbb{R},c_j>0\right\}.$$

For $f\in\mathcal{A}$, it follows immediately from (6.85) that

$$\lim_{t\to 0^+}\int_0^\infty f\,d\nu_t=\int_0^\infty f\,d\nu_0. \tag{6.86}$$

Let $C_0(\overline{\mathbb{R}}_+,\mathbb{R})$ denote the set of continuous real-valued functions on $[0,\infty)$ that vanish at ∞. By the Stone–Weierstrass theorem [77, Thm. 7.32], $\mathcal{A}$ is dense in $C_0(\overline{\mathbb{R}}_+,\mathbb{R})$ with respect to the uniform topology. The measures ν_t are uniformly bounded, since $\lim_{t\to 0^+}\nu_t[0,\infty)=\nu_0[0,\infty)=A$. Therefore (6.86) holds for all $f\in C_0(\overline{\mathbb{R}}_+,\mathbb{R})$.

Now for $\delta>0$, let

$$f_\delta(x)\begin{cases}1, & x\in[0,1],\\ 1-(x-1)/\delta, & x\in(1,1+\delta),\\ 0, & x\geq 1.\end{cases}$$

Then

$$\int_0^1 e^x\,d\mu_t(x)\leq\int_0^\infty e^x f_\delta(x)\,d\nu_t(x),$$

and taking the limit of both sides using (6.86) gives

$$\begin{aligned}\limsup_{t\to 0^+}\int_0^1 e^x d\mu_t &\leq \int_0^\infty e^x f_\delta(x)\,d\nu_0(x)\\ &\leq \int_0^1 e^x\,d\nu_0(x)+C\delta.\end{aligned}$$

Hence,

$$\limsup_{t\to 0^+}\int_0^1 e^x d\mu_t\leq\int_0^1 e^x\,d\nu_0.$$

Combining this with the corresponding continuous approximation of $\chi_{[0,1]}$ from below yields

$$\lim_{t\to 0^+}\int_0^1 e^x\, d\nu_t(x) = \int_0^1 e^x\, d\nu_0(x) \tag{6.87}$$

This is the claimed result, since

$$\mu[0,s] = s^\alpha \int_0^1 e^x\, d\nu_{1/s}(x),$$

and

$$\int_0^1 e^x\, d\nu_0(x) = \frac{A}{\Gamma(\alpha+1)}.$$

□

With Karamata's Tauberian theorem, we now have all the ingredients for the proof of the Weyl asymptotic.

Proof of Theorem 6.27 For $m > n/2$, let

$$g(t) := \sum_{k=1}^{\infty} \frac{1}{(\lambda_k + t)^m}.$$

We can write this in terms of the counting function $N_\Omega(t)$ as a Stieljes integral,

$$g(t) = \int_0^\infty \frac{1}{(x+t)^m} dN_\Omega(x). \tag{6.88}$$

Setting $h = t^{-1/2}$ in Theorem 6.32 gives the asymptotic

$$g(t) \sim A\frac{\Gamma(m-\frac{n}{2})}{\Gamma(m)} t^{-m+\frac{n}{2}} \tag{6.89}$$

as $t \to \infty$, where

$$A := (4\pi)^{-\frac{n}{2}} \operatorname{vol}(\Omega).$$

For the integrand in (6.88), we can set

$$\frac{1}{(x+t)^m} = \frac{1}{\Gamma(m)}\int_0^\infty y^{m-1} e^{-(t+x)y}\, dy.$$

After switching the order of integration using Fubini's theorem, this gives

$$g(t) = \frac{1}{\Gamma(m)} \int_0^\infty y^{m-1} e^{-ty} f(y)\, dy, \tag{6.90}$$

where

$$f(y) := \int_0^\infty e^{-yx} dN_\Omega(x).$$

Theorem 6.33 (Karamata) now applies to the integral (6.90). From the asymptotic (6.89) we deduce that

$$\int_0^s y^{m-1} f(y)\, dy \sim \frac{A}{(m-\frac{n}{2})} s^{m-\frac{n}{2}}, \tag{6.91}$$

as $s \to 0^+$.

To proceed, we would like to differentiate (6.91) with respect to s. Although differentiation of asymptotic limits is not generally possible, we have an advantage here in the fact that $f(y)$ is monotonically decreasing. For $h \in (0, s)$ this gives

$$\int_{s-h}^s y^{m-1} f(y)\, dy \geq h(s-h)^{m-1} f(s),$$

or

$$f(s) \leq \frac{1}{h(s-h)^{m-1}} \int_{s-h}^s y^{m-1} f(y)\, dy. \tag{6.92}$$

For convenience, set $\nu := m - n/2$, so that the right-hand side of (6.91) is $a\nu^{-1}s^\nu$. For $\varepsilon > 0$ and $u \leq c_\varepsilon$,

$$\left| u^{-\nu} \int_0^u y^{m-1} f(y)\, dy - a\nu^{-1} \right| < \varepsilon.$$

Applying this to the integral in (6.92) yields

$$f(s) \leq \frac{A\nu^{-1}[s^\nu - (s-h)^\nu] + 2\varepsilon s^\nu}{h(s-h)^{m-1}}.$$

for $0 < h < s \leq c_\varepsilon$. Setting $h = \sqrt{\varepsilon} s$ then gives

$$\frac{f(s)}{s^{\nu-m}} \leq \frac{A\nu^{-1}[1-(1-\sqrt{\varepsilon})^\nu] + 2\varepsilon}{\sqrt{\varepsilon}(1-\sqrt{\varepsilon})^{m-1}}.$$

As $s \to 0$ we can also take $\varepsilon \to 0$, so that

$$\limsup_{s\to 0^+} \frac{f(s)}{s^{\nu-m}} \le \lim_{\varepsilon\to 0^+} \frac{A\nu^{-1}[1-(1-\sqrt{\varepsilon})^\nu]+2\varepsilon}{\sqrt{\varepsilon}(1-\sqrt{\varepsilon})^{m-1}}$$
$$= A.$$

As similar argument based on the integral from s to $s+h$ shows that

$$\liminf_{s\to 0^+} \frac{f(s)}{s^{\nu-m}} \ge A.$$

Therefore, $f(s) \sim As^{\nu-m}$ as $s \to 0^+$, which translates to

$$f(s) \sim As^{-\frac{n}{2}}. \tag{6.93}$$

Applying Karamata's theorem to $f(s)$ yields

$$N_\Omega(x) \sim \frac{A}{\Gamma(\frac{n}{2}+1)} x^{\frac{n}{2}},$$

as $x \to \infty$. Since

$$\frac{A}{\Gamma(\frac{n}{2}+1)} = (2\pi)^{-n}\omega_n \operatorname{vol}(\Omega),$$

this completes the proof. □

6.6 Nodal Domains

The *nodes* of a vibrating string are the positions on the string that remain at rest during the vibration. For a "pure tone" solution which exhibits a single frequency, the spatial component is an eigenfunction, and the nodes correspond to its zeros, as illustrated in Figure 6.11. Since the kth eigenfunction is proportional to $\sin(\pi kx/\ell)$, it has exactly $k-1$ nodes. This property holds for more general one-dimensional eigenvalue problems, by the Sturm–Liouville theory.

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set, with Dirichlet eigenvalues $\{\lambda_k\}$ written in increasing order as above. As noted earlier, we can assume that the eigenfunctions

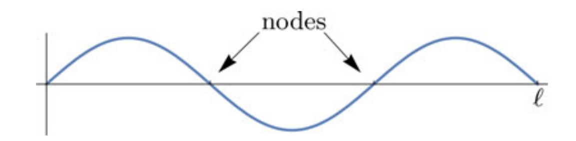


Fig. 6.11 A string eigenfunction with two nodes

ψ_k are real-valued, by separating the real and imaginary parts if necessary. The *nodal domains* of ψ_k are defined to be the connected components of the open set $\{\psi_k \neq 0\} \subset \Omega$. Figure 6.12 shows a sample nodal domain plot for a triangle.

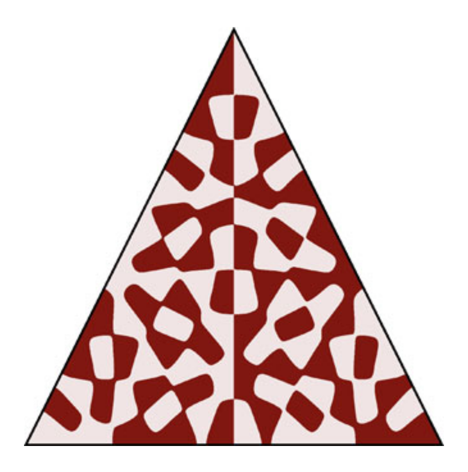


Fig. 6.12 The 128th eigenfunction of the equilateral triangle has 30 nodal domains

We would naturally expect the nodal pattern to become more intricate as the frequency increases. The relationship can be quantified by estimating the maximum possible number of domains for each eigenfunction. We first consider the case of the eigenfunction ψ_1 corresponding to the bottom of the spectrum.

Theorem 6.34. *Let $\Omega \subset \mathbb{R}^n$ be a bounded, connected open set. The lowest Dirichlet eigenvalue λ_1 is simple, and its eigenfunction has a single nodal domain.*

Proof Let $\{\psi_k\}$ be the orthonormal basis for $L^2(\Omega)$ consisting of Dirichlet eigenfunctions. We can split ψ_1 into positive and negative components by setting

$$\psi_1^{\pm}(x) := \max\{\psi_1(x), 0\}.$$

To check that these components lie in $H_0^1(\Omega)$, let

$$F_\varepsilon(t) := \begin{cases} te^{-\varepsilon^2/(t-\varepsilon)^2}, & t > \varepsilon, \\ 0, & t \leq \varepsilon. \end{cases}$$

Then $F_\varepsilon \circ \psi_1$ is smooth, by the interior regularity of ψ_1 (Theorem 6.15), and has support equal to $\{\psi_1 \geq \varepsilon\}$, which is compact. It is easy to check that $F_\varepsilon \circ \psi_1 \to \psi_1^+$ as $\varepsilon \to 0$, with respect to the H^1 norm, proving that $\psi_1^+ \in H_0^1(\Omega)$. A similar argument applies to ψ_1^-.

Since the supports of ψ_1^+ and ψ_1^- intersect in a set of measure zero, both the functions and their derivatives are orthogonal. In particular

$$\lambda_1 = \|\nabla\psi_1\|^2 = \|\nabla\psi_1^+\|^2 + \|\nabla\psi_1^-\|^2. \tag{6.94}$$

On the other hand, we have

$$\|\nabla\psi_1^\pm\|^2 \geq \lambda_1\|\psi_1^\pm\|^2$$

by (6.20), and $\|\psi_1^+\|^2 + \|\psi_1^-\|^2 = 1$. In view of (6.94), this shows that

$$\|\nabla\psi_1^\pm\|^2 = \lambda_1\|\psi_1^\pm\|^2. \tag{6.95}$$

Since ψ_1 is nonzero, at least one of the components $\psi_1^\pm$ must be nonzero. After adjusting the sign, if needed, let us assume that ψ_1^+ is nonzero. If the basis expansion of ψ_1^+ is denoted by $\sum c_k\psi_k$, then (6.95) implies that

$$\sum_{k=1}^\infty \lambda_k|c_k|^2 = \lambda_1\sum_{k=1}^\infty |c_k|^2.$$

Since $\lambda_k \geq \lambda_1$, we conclude that $c_k = 0$ unless $\lambda_k = \lambda_1$. In other words, ψ_1^+ is a linear combination of eigenfunctions with eigenvalue λ_1. Therefore, ψ_1^+ is itself an eigenfunction. Unique continuation (Corollary 6.16) then implies that ψ_1^+ cannot vanish on a nonempty open set. Therefore supp $\psi_1^+ = \overline{\Omega}$ and $\psi_1 \geq 0$.

To prove the simplicity of the first eigenvalue, suppose that $\lambda_1 = \lambda_2$. The argument above then applies to ψ_2 also, which means that $\psi_2 \geq 0$ in Ω. This is a contradiction, because ψ_1 and ψ_2 are orthogonal. Therefore $\lambda_2 > \lambda_1$. □

Following Theorem 6.34, we adopt the standard convention that $\psi_1 \geq 0$. This inequality can be sharpened in the interior by noting that $\psi_1 \geq 0$ implies that $-\Delta\psi_1 = \lambda_1\psi_1 \geq 0$, i.e., ψ_1 is *superharmonic*. Superharmonic functions satisfy a strict minimum principle: a local minimum cannot occur at an interior point unless the function is constant. (See, e.g., [13, Thm. 9.5].) Thus $\psi_1 > 0$ on Ω.

In 1923 Richard Courant used an extension of the argument we have given for ψ_1 to prove the following bound on the number of nodal domains for higher eigenfunctions.

Theorem 6.35 (Courant). *The Dirichlet eigenfunction ψ_k has at most k nodal domains.*

Proof Suppose that ψ_k has at least k nodal domains. We label these as $V_1, \ldots, V_k$, and define the restrictions

$$\psi_k^{(j)}(x) := \begin{cases} \psi_k(x), & x \in V_j \\ 0, & x \notin V_j, \end{cases}$$

for $j = 1, \ldots, k$. Note that each $\psi_k^{(j)} \in H_0^1(\Omega)$ by the argument used in the proof of Theorem 6.34.

Counting dimensions shows that there exists a nonzero element $u \in \operatorname{span}\{\psi_k^{(1)}, \dots, \psi_k^{(k)}\}$ such that u is orthogonal to each ψ_i for $i = 1, \dots, k-1$. The eigenfunction basis decomposition of u thus has the form

$$u = \sum_{i=k}^{\infty} \langle \psi_i, u \rangle \psi_i.$$

This implies that

$$\|\nabla u\|^2 = \sum_{i=k}^{\infty} \lambda_i |\langle \psi_i, u \rangle|^2. \tag{6.96}$$

On the other hand, since u is a linear combination of disjoint components of ψ_k, we can argue as in Theorem 6.34 that

$$\|\nabla u\|^2 = \lambda_k \|u\|^2 = \lambda_k \sum_{i=k}^{\infty} |\langle \psi_i, u \rangle|^2.$$

Comparing this to (6.96) shows that $\langle \psi_i, u \rangle = 0$ unless $\lambda_i = \lambda_k$, implying that u is itself eigenfunction with eigenvalue λ_k.

By construction, u vanishes outside $V_1 \cup \cdots \cup V_k$, but unique continuation (Corollary 6.16) implies that u cannot vanish on an open set. It follows that ψ_k cannot have more than k nodal domains. □

6.7 Isoperimetric Inequalities and Minimal Eigenvalues

In 1894 the physicist Lord Rayleigh conjectured that among all vibrating membranes of a given area, the lowest fundamental tone is attained only if the membrane is circular. In other words, among all bounded open sets $\Omega \subset \mathbb{R}^2$ of a fixed area, the disk is the unique minimizer of $\lambda_1(\Omega)$. From the calculation in Example 6.2, we can see that for a disk in $\mathbb{R}^2$ of radius r, the lowest eigenvalue is

$$\lambda_1(\mathbb{D}) = j_{0,1}^2,$$

where $j_{0,1} \approx 2.4048$ is the first zero of the Bessel function J_0. By scaling, we can thus state Rayleigh's conjecture in the form

$$\lambda_1(\Omega) \geq \frac{\pi j_{0,1}^2}{\operatorname{vol}(\Omega)}.$$

This result was proven independently by Georg Faber and Edgar Krahn in the 1920s.

Theorem 6.36 (Faber–Krahn Inequality). *Suppose that Ω is a bounded open set $\mathbb{R}^n$, and let $B \subset \mathbb{R}^n$ be a ball with* $\mathrm{vol}(B) = \mathrm{vol}(\Omega)$. *The lowest Dirichlet eigenvalues satisfy,*

$$\lambda_1(\Omega) \geq \lambda_1(B), \tag{6.97}$$

with equality only if Ω is a ball.

We can make the inequality in Theorem 6.36 more explicit by computing the lowest eigenvalue of the unit ball $\mathbb{B}^n \subset \mathbb{R}^n$. A straightforward generalization of the calculation from Example 6.2 gives

$$\lambda_1(\mathbb{B}) = j^2_{\frac{n}{2}-1,1},$$

where $j_{\frac{n}{2}-1,1}$ denotes the first zero of the Bessel function $J_{\frac{n}{2}-1}$. Using scaling and the fact that $\mathrm{vol}(\mathbb{B}) = \pi^{\frac{n}{2}}/\Gamma(\frac{n}{2}+1)$, the inequality (6.97) can be written as

$$\lambda_1(\Omega) \geq \mathrm{vol}(\Omega)^{-\frac{2}{n}} \frac{\pi j^2_{\frac{n}{2}-1,1}}{\Gamma(\frac{n}{2}+1)^{\frac{2}{n}}}.$$

The Faber–Krahn inequality is a consequence of the classical isoperimetric inequality. Suppose U is a bounded open subset of $\mathbb{R}^n$ with C^1 boundary. If B is a ball with the same volume as U, then

$$\mathrm{area}(\partial U) \geq \mathrm{area}(\partial B), \tag{6.98}$$

with equality only if U is a ball. Here "area" refers the $(n-1)$-dimensional hypersurface area in $\mathbb{R}^n$. For the proof of (6.98), see, for example, Chavel [20, Thm. III.2.3 & III.2.4].

The derivation of Theorem 6.36 from the isoperimetric inequality relies on a change-of-variables formula for the decomposition of an integral in terms of level surfaces, called the *co-area formula*. Although the co-area formula can be generalized to include functions with minimal regularity, we will limit our attention to the relatively simple smooth case, following a proof given in Chavel [19, §IV.1].

Lemma 6.37 (Co-area Formula). *Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded open set, and $f : \Omega \to \mathbb{R}$ a smooth function. Let $I \subset \mathbb{R}$ be a bounded interval that contains only regular values of f. For $g \in C(\Omega)$,*

$$\int_{f^{-1}(I)} g\, d^n x = \int_I \int_{\{f=t\}} \frac{g}{|\nabla f|}\, dS_t\, dt, \tag{6.99}$$

where dS_t denotes the surface area element on $\{f = t\}$.

Proof It suffices to prove the integration formula in the neighborhood of a point a such that $f(a) = t_0 \in I$. Let $\sigma : U \subset \mathbb{R}^{n-1} \to \mathbb{R}^n$ be a regular parametrization of a neighborhood of a in the level surface $\{f = f(a)\}$. We can define a

local parametrization of the level surfaces by integrating the normal vector field $\nabla f/|\nabla f|^2$, starting points in the patch $\sigma(U)$. That is, for each $y \in U$, let $\Phi(y, t)$ be the trajectory in $\mathbb{R}^n$ such that

$$\Phi(y, t_0) = \sigma(y), \quad \partial_t \Phi(y, t) = \frac{\nabla f}{|\nabla f|^2}(\Phi(y, t)). \tag{6.100}$$

Existence and uniqueness of solutions of (6.100) are guaranteed for $|t - t_0| < \delta$ with $\delta > 0$ sufficiently small, by the Picard–Lindelöf theorem from ODE theory (see Theorem 9.8). Moreover, Φ is smooth and defines a local coordinate system for $\mathbb{R}^n$ in some neighborhood of a.

By the chain rule,

$$\begin{aligned} \partial_t f(\Phi(y, t)) &= \nabla f(\Phi(y, t)) \cdot \partial_t \Phi(y, t) \\ &= 1. \end{aligned}$$

Since $f(\Phi(y, t_0)) = t_0$, this shows that

$$f(\Phi(y, t)) = t,$$

for all $|t - t_0| < \delta$.

By the standard Jacobian formula, the volume form on $\mathbb{R}^n$ translates to the coordinates (y, t) as

$$d^n x = \left| \det\left[\frac{\partial \Phi}{\partial y_1}, \dots, \frac{\partial \Phi}{\partial y_{n-1}}, \frac{\partial \Phi}{\partial t} \right] \right| d^{n-1}y\, dt. \tag{6.101}$$

On the other hand, under the parametrization $y \mapsto \Phi(y, t)$, the surface area element on $\{f = t\}$ is given by

$$dS_t = \left| \det\left[\frac{\partial \Phi}{\partial y_1}, \dots, \frac{\partial \Phi}{\partial y_{n-1}}, \nu \right] \right| d^{n-1}y, \tag{6.102}$$

where ν denotes the unit normal. Since $\nu = \nabla f/|\nabla f|$, we have (6.100), and we have

$$\nu = |\nabla f| \, \frac{\partial \Phi}{\partial t}.$$

A comparison of (6.101) and (6.102) shows that

$$d^n x = \frac{1}{|\nabla f|} dS_t\, dt,$$

within the coordinate patch covered by Φ. This local argument applies wherever f is regular. □

With the co-area formula, and assuming the isoperimetric inequality, we are now prepared to prove the Faber–Krahn inequality.

Proof of Theorem 6.36 Let ϕ_1 be an eigenfunction of Ω with eigenvalue λ_1. The strategy for the proof is to compare the Rayleigh quotients of ϕ_1 with those of a radial function ψ called the *symmetric decreasing rearrangement* of ϕ_1.

By Theorem 6.15, ϕ_1 is an analytic function, and by Theorem 6.34 we can assume that $\phi_1 \geq 0$ within Ω. Let $T = \max_\Omega \phi_1$, and define $\mathcal{R} \subset [0, T]$ to be the set of regular values of ϕ_1. By Sard's theorem, the set of critical values $[0, T]\backslash\mathcal{R}$ has measure zero.

Consider the decreasing function V on $[0, T]$ defined by

$$V(t) := \mathrm{vol}\{x \in \Omega : \phi_1(x) > t\}.$$

By the inverse function theorem and the smoothness of ϕ_1, $V(t)$ is smooth for all $t \in \mathcal{R}$. The analyticity of ϕ_1 implies that critical points are contained in submanifolds of lower dimension. In particular, the set of critical points of ϕ_1 has measure zero, and hence V is continuous on $[0, T]$.

For each $t \in [0, 1]$, let $r(t)$ be the radius of a ball with volume equal to $V(t)$, i.e.,

$$r(t) := \left(\frac{V(t)}{\mathrm{vol}(\mathbb{B})}\right)^{1/n}.$$

By the properties of $V(t)$, $r(t)$ is decreasing and continuous on $[0, T]$, and smooth on $\mathcal{R}$. Thus $r(t)$ admits an inverse function $\psi : [0, r_0] \to [0, T]$. We can interpret $\psi(r)$ as a radial function defined on the ball $B(0, r_0)$. By construction we have that

$$\mathrm{vol}\{\psi > t\} = V(t). \tag{6.103}$$

This is the defining condition for the symmetric decreasing rearrangement.

The next step in the proof is to check that symmetric rearrangement preserves the L^2 norm. First, note that since $[0, T]\backslash\mathcal{R}$ has measure zero, the co-area formula (6.99) implies that

$$V(t) = \int_t^T \int_{\{\phi_1 = t\}} \frac{dS}{|\nabla\phi_1|}\, dt.$$

Therefore,

$$V'(t) = -\int_{\{\phi_1 = t\}} \frac{dS}{|\nabla\phi_1|}. \tag{6.104}$$

We can thus compute that

$$\begin{aligned}
\|\phi_1\|^2 &= \int_\Omega \phi_1^2\, d^n x \\
&= \int_0^T \int_{\{\phi_1 = t\}} t^2 \frac{dS}{|\nabla\phi_1|}\, dt \\
&= -\int_0^T t^2 V'(t)\, dt.
\end{aligned}$$

Since this equation involves only $V(t)$, we deduce from (6.103) that

$$\|\psi\|^2 = \|\phi_1\|^2. \tag{6.105}$$

The next step is to compare $\|\nabla\phi_1\|$ with $\|\nabla\psi\|$. By the co-area formula,

$$\|\nabla\phi_1\|^2 = \int_0^T \int_{\{\phi_1=t\}} |\nabla\phi_1|\, dS\, dt, \tag{6.106}$$

with a similar expression for $\|\nabla\psi\|^2$.

For $t \in \mathcal{R}$, the inner integral can be estimated in terms of the area of the level set by an application of the Cauchy–Schwarz inequality,

$$\begin{aligned}\text{area}(\{\phi_1 = t\})^2 &= \left(\int_{\{\phi_1=t\}} dS\right)^2 \\ &\le \left(\int_{\{\phi_1=t\}} |\nabla\phi_1|\, dS\right)\left(\int_{\{\phi_1=t\}} \frac{dS}{|\nabla\phi_1|}\right),\end{aligned}$$

In view of (6.104), this gives the estimate

$$\int_{\{\phi_1=t\}} |\nabla\phi_1|\, dS \ge \frac{(\text{area}\{\phi_1 = t\})^2}{-V'(t)}. \tag{6.107}$$

Since $\text{vol}(\{\phi_1 > t\}) = \text{vol}(\{\psi > t\})$, by construction, and $\{\psi_1 = 0\}$ is smooth for $t \in \mathcal{R}$, the isoperimetric inequality (6.98) implies that

$$\text{area}\{\phi_1 = t\} \ge \text{area}\{\psi = t\}, \tag{6.108}$$

with equality only if $\{\phi_1 = t\}$ is a sphere. Note also that, since ψ is radial,

$$\begin{aligned}\text{area}(\{\psi = t\})^2 &= \left(\int_{\{\psi=t\}} |\nabla\psi|\, dS\right)\left(\int_{\{\psi=t\}} \frac{dS}{|\nabla\psi|}\right) \\ &= -V'(t)\left(\int_{\{\psi=t\}} |\nabla\psi|\, dS\right).\end{aligned}$$

Thus, by (6.107) and (6.108),

$$\int_{\{\phi_1=t\}} |\nabla\phi_1|\, dS \ge \int_{\{\psi=t\}} |\nabla\psi|\, dS \tag{6.109}$$

for $t \in \mathcal{R}$, with equality only if $\{\phi_1 = t\}$ is a sphere.

Applying (6.109) to (6.106) now shows that

$$\|\nabla\phi_1\|^2 \geq \|\nabla\psi\|^2, \tag{6.110}$$

with equality only if $\{\phi_1 = t\}$ is spherical for all $t \in \mathcal{R}$. Equality thus occurs only if ϕ_1 is radial, implying that Ω is a ball. □

A unique symmetric decreasing rearrangement can be constructed for any measurable function $\mathbb{R}^n \to \mathbb{R}$. The bound (6.110), which is the essential point in our proof of Faber–Krahn, is called the Pólya–Szegő inequality [68]. It can be proven directly, without relying on an isoperimetric inequality; see, e.g., Lieb and Loss [59, §7.17].

The Faber–Krahn inequality is part of a family of isoperimetric results for eigenvalues in different contexts. For example, the Szegő–Weinberger inequality [88, 95] gives an analog of Faber–Krahn for Neumann eigenvalues. Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with smooth boundary $\partial\Omega$. If B is a ball with the same volume as Ω, then

$$\mu_2(\Omega) \leq \mu_2(B),$$

with equality only is Ω is a ball.

Another famous result, originally conjectured by Payne, Pólya, and Weinberger [65], was proven in 1990 by Ashbaugh and Benguria [5]. This says that the ratio of the first two Dirichlet eigenvalues, λ_2/λ_1, is minimized for the sphere. In particular, to determine whether a drum is circular, it is enough to hear the first two eigenfrequencies.

6.8 Exercises

6.1. On $L^2(0, 1)$, let A be the self-adjoint extension of the Laplacian $-\partial^2$ with the Robin boundary conditions,

$$f'(0) = \alpha f(0), \quad f'(1) = -\alpha f(1),$$

for a constant $\alpha \in \mathbb{R}$, defined as in Example 3.32.

(a) Show that A is positive for $\alpha \geq 0$ and semi-bounded, as defined in (3.25), for $\alpha < 0$. [Hint: To estimate $|f(0)|^2$ and $|f(1)|^2$ in terms of $\|f\|$ and $\|f'\|$, integrate the derivative of $(2t-1)|f(t)|^2$.]
(b) Prove that A has discrete spectrum.

(c) Let $\lambda_1 \leq \lambda_2 \leq \dots$ be the eigenvalues of A. Prove that

$$\lambda_k = \min_{W \in \Lambda_k} \left\{ \max_{u \in W, \|u\|=1} Q_\alpha[u] \right\},$$

where Λ_k denotes the set of k-dimensional subspaces of $H^1(0, 1)$ and

$$Q_\alpha[u] := \|u'\|^2 + \alpha\Big(|u(0)|^2 + |u(1)|^2\Big)$$

(d) For $\alpha \geq 0$, use the min–max principle to prove that

$$\pi^2(k-1)^2 \leq \lambda_k \leq \pi^2 k^2.$$

6.2. Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded open set, and that $\{\Omega_j\}$ is a sequence of open sets such that

$$\Omega_1 \subset \Omega_2 \subset \dots,$$

with $\Omega = \cup \Omega_j$. Let $\lambda_k(\Omega_j)$ denote the kth Dirichlet eigenvalue of Ω_j, and similarly for Ω. Use the min–max principle to show that, for each k,

$$\lim_{j \to \infty} \lambda_k(\Omega_j) = \lambda_k(\Omega).$$

6.3. Let λ_1 be the lowest Dirichlet eigenvalue of a bounded open set $\Omega \subset \mathbb{R}^n$. For a finite-dimensional subspace $\mathcal{A} \subset H_0^1(\Omega)$, define β_1 to be the Rayleigh–Ritz approximation to λ_1 defined by (6.46). Suppose that the orthogonal projection Q onto $\mathcal{A}^\perp$ is bounded as an operator on $H_0^1(\Omega)$. Estimate $\beta_1 - \lambda_1$ in terms of the operator norms of Q acting on $L^2(\Omega)$ and $H_0^1(\Omega)$.

6.4. On an open set $\Omega \subset \mathbb{R}^n$, let L be the differential operator,

$$Lu := -\sum_{i,j=1}^{n} \partial_i \big[a_{ij}(x)\, \partial_j u\big], \tag{6.111}$$

where the coefficients $a_{ij} \in C^\infty(\overline{\Omega}; \mathbb{R})$ satisfy $a_{ij} = a_{ji}$ for all i, j. Assume that L is *uniformly elliptic* on Ω, which means that there exists $c > 0$ such that

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq c|\xi|^2, \tag{6.112}$$

for all $x \in \Omega$ and $\xi \in \mathbb{R}^n$.

(a) To define the Dirichlet self-adjoint extension of L to $L^2(\Omega)$, set

$$\mathcal{D}(L) := \left\{u \in H_0^1(\Omega) : Lu \in L^2(\Omega)\right\},$$

where Lu is defined in the weak sense. Prove that L is a positive self-adjoint operator on this domain.

(b) If Ω is bounded, prove that L has discrete spectrum with strictly positive eigenvalues.

6.5. Prove Carleman's local version of the Weyl asymptotic: The Dirichlet eigenfunctions $\{\psi_k\}$ of a bounded open set $\Omega \subset \mathbb{R}^n$ satisfy

$$\lim_{m\to\infty} \frac{1}{m}\sum_{k=1}^{m} \psi_k(x)\overline{\psi_k(y)} = \begin{cases} \dfrac{1}{\mathrm{vol}(\Omega)}, & x = y, \\ 0, & x \neq y. \end{cases}$$

6.6. The wave equation for vibrations of a metal plate involves the *biharmonic* operator (or "bi-Laplacian") Δ^2. For a bounded open set $\Omega \subset \mathbb{R}^n$, let A denote the Friedrichs extension of Δ^2 acting on $C_0^\infty(\Omega)$. Prove that A has purely discrete spectrum by showing that $\mathcal{D}(A) \subset H_0^1(\Omega)$. [Hint: You can estimate the H^1 norm in terms of the quadratic form $Q[u] := \|\Delta u\|^2$ by exploiting the positivity of $(\Delta - 1)^2$.]

6.7. On the interval $[0, 1]$, define the differential operator

$$Lu := -(\rho u')',$$

where $\rho \in C^\infty[0, 1]$ is strictly positive. Define the Dirichlet realization of L on $L^2[0, 1]$ as in Exercise 6.4. Let $\{\lambda_k\}$ be the eigenvalues of L, in ascending order, with $\{\phi_k\}$ the corresponding real-valued eigenfunctions. You may assume that the eigenfunctions are contained in $C^\infty[0, 1]$, which is a consequence of elliptic regularity.

(a) Prove that the eigenvalues satisfy strict inequalities,

$$0 < \lambda_1 < \lambda_2 < \dots.$$

(b) For each eigenfunction ϕ, consider $t \mapsto (\rho\phi', \phi)$ as a parametrized curve in $\mathbb{R}^2$. We can assume that $\phi'(0) > 0$, so that the curve starts on the positive real axis. Let η denote the polar-coordinate angle of $(\rho\phi', \phi)$, traced continuously starting from $\eta(0) = 0$. Show that η is a smooth, strictly increasing function on $[0, 1]$.

(c) Let η_k denote the angle function associated with the eigenfunction ϕ_k, defined as in (b). Show that $\eta_{k+1}(t) > \eta_k(t)$ for all $t > 0$.

(d) Use (c) to prove that ϕ_k has exactly k nodal domains.

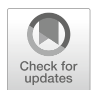
Notes

For more background on the Friedrichs extension method used in Section 6.1, see Davies [24, §4.4], Reed and Simon [70, §X.3], or Schmüdgen [80, Chapter 10]. The most general versions of the Sobolev space results discussed in Section 6.2 are developed in Gilbarg and Trudinger [36, Chapter 7].

Weyl's law has been adapted to many other settings and sharpened in various ways. For more on this extensive history, see the survey articles by Arendt et al. [1] or Ivrii [46]

The Faber–Krahn inequality is covered in detail in [20]. For background on related inequalities, see the survey article by Ashbaugh [4].

Chapter 7
Schrödinger Operators



In quantum mechanics, the motion of a single particle in $\mathbb{R}^n$ is described by a normalized state function $\psi \in L^2(\mathbb{R}^n)$, such that $|\psi|^2$ is the probability density for the particle's location. Observable quantities such as energy and momentum are represented by (unbounded) self-adjoint operators on $L^2(\mathbb{R}^n)$. For example, the classical coordinate x_j is represented by the multiplication operator M_{x_j}, and the momentum component p_j by the differential $-i\hbar\partial_{x_j}$, where $\hbar$ is Planck's constant. The possible values of an observable are given by the spectrum of the associated operator, and the distribution of these values for a particular quantum state ψ corresponds to its spectral decomposition.

For a classical particle of mass m the kinetic energy is $p^2/2m$. If the potential energy is represented by a real-valued function $V(x)$, the classical total energy (called the *Hamiltonian* function) is

$$E := \frac{p^2}{2m} + V(x).$$

The quantum version of the Hamiltonian, according to the prescription described above, is the operator

$$H := -\frac{\hbar^2}{2m}\Delta + V \tag{7.1}$$

(with the potential acting by multiplication).

An operator of the form (7.1) is called a *Schrödinger operator*. In 1926, Erwin Schrödinger applied the quantization scheme described above to the case of the electron in a hydrogen atom, where $V(x)$ is the Coulomb potential for the electric field generated by a single proton, assumed to be fixed at the origin. In Gaussian units this potential is given by

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1_7

$$V(x) := -\frac{q^2}{r}, \tag{7.2}$$

where q is the elementary charge, and $r := |x|$. Schrödinger showed that the spectrum of the associated Hamiltonian has discrete spectrum below zero. Moreover, the eigenvalues correspond precisely to the energy levels that Niels Bohr had postulated for atomic hydrogen in 1913 to explain absorption and emission lines of hydrogen gas observed in the nineteenth century. This derivation of the empirical result from the spectrum of H was a major achievement for the new quantum mechanical theory.

For mathematical purposes, we usually scale out the physical constants and write a Schrödinger operator as

$$H = -\Delta + V.$$

In this chapter we will develop the spectral theory of Schrödinger operators for various classes of real-valued potentials on $\mathbb{R}^n$. Although the case $n = 3$ is of special importance, there are physical situations, usually involving crystal lattices, where particle motion is limited to one or two dimensions. Moreover, systems with multiple interacting particles are described by Schrödinger operators on higher dimensional spaces.

If we assume that the potential V is real-valued and locally L^2, then H is well defined as a symmetric operator on $C_0^\infty(\mathbb{R}^n)$. As we have seen in Examples 3.23 and 3.26, the Laplacian itself is essentially self-adjoint on $C_0^\infty(\mathbb{R}^n)$, with the domain of the self-adjoint extension given by $H^2(\mathbb{R}^n)$. Thus, under the additional assumption that V is bounded, Lemma 3.27 implies that H is essentially self-adjoint on $C_0^\infty(\mathbb{R}^n)$ and self-adjoint on $H^2(\mathbb{R}^n)$.

These assumptions on the potential are too restrictive for many problems in quantum mechanics. The Coulomb potential (7.2), for example, is singular and unbounded. Establishing self-adjoint extensions for a broad class of potentials that includes the important physical cases is a fundamental goal in the theory of Schrödinger operators. If we can show that a quantum Hamiltonian operator is essentially self-adjoint on $C_0^\infty(\mathbb{R}^n)$, then its spectrum is canonically defined and not dependent on other choices.

7.1 Positive Potentials

Because the Laplacian is a positive operator, it is possible to establish the essential self-adjointness of $-\Delta + V$ under fairly general conditions when $V \geq 0$. In fact, since shifting V by a constant does not affect the extension properties, the results that we develop for the positive case extend immediately to potentials which are *semi-bounded*, meaning pointwise bounded from below. This shift is a trivial adjustment, so in this section we will restrict our attention to the case $V \geq 0$ for convenience.

7.1.1 Essential Self-adjointness

By the criteria given in Theorem 3.30 and Exercise 3.11, the essential self-adjointness of $-\Delta + V$ on $C_0^\infty(\mathbb{R}^n)$ can be proven for $V \geq 0$ by establishing that $(-\Delta + V)^* + 1$ is injective. As we will see below, this amounts to ruling out any possible weak L^2 solutions of

$$(-\Delta + V + 1)u = 0. \tag{7.3}$$

We first consider the relatively straightforward case where V is locally bounded.

Theorem 7.1. *Suppose that $V \in L^\infty_{\mathrm{loc}}(\mathbb{R}^n)$ is real-valued with $V \geq 0$. Then $-\Delta+V$ is essentially self-adjoint on $C_0^\infty(\mathbb{R}^n)$.*

Proof Let $A = -\Delta + V$ with the domain $\mathcal{D}(A) = C_0^\infty(\mathbb{R}^n)$, and consider $u \in \mathcal{D}(A^*)$. In order to apply results from vector calculus, we want to show that u is at least locally H^2.

For $\psi \in C_0^\infty(\mathbb{R}^n)$, it is easy to see from the definition of the adjoint domain that $\psi u \in \mathcal{D}(A^*)$ also. This means that $(-\Delta + V)\psi u \in L^2(\mathbb{R}^n)$, with the derivative defined in the weak sense. Since $V\psi u \in L^2$ by the assumption that V is locally bounded, we conclude that $-\Delta(\psi u) \in L^2(\mathbb{R}^n)$. Taking the Fourier transform implies that $|\xi|^2\widehat{\psi u} \in L^2(\mathbb{R}^n)$, which gives $\psi u \in H^2(\mathbb{R}^n)$ by (2.26). Since ψ was arbitrary, this argument shows that

$$\mathcal{D}(A^*) \subset H^2_{\mathrm{loc}}(\mathbb{R}^n).$$

Our main goal is to show that $A^* + 1$ is injective. For $u \in \mathcal{D}(A^*)$ assume for the sake of contradiction that

$$(A^* + 1)u = 0. \tag{7.4}$$

This is equivalent to the statement that

$$\langle u, (-\Delta + V + 1)\phi \rangle = 0 \tag{7.5}$$

for all $\phi \in C_0^\infty(\mathbb{R}^n)$. In other words, u is a weak solution of (7.3). The left-hand side of (7.5) is continuous as a function of ϕ with respect to the H^2 topology. Therefore, since $u \in H^2_{\mathrm{loc}}(\mathbb{R}^n)$, for $\chi \in C_0^\infty(\mathbb{R}^n)$ we can use approximation of $\chi^2 u$ by functions in $C_0^\infty(\mathbb{R}^n)$ to show that

$$\left\langle u, (-\Delta + V + 1)\chi^2 u \right\rangle = 0.$$

Since $V \geq 0$, this implies

$$\|\chi u\|^2 \leq \left\langle u, \Delta(\chi^2 u) \right\rangle. \tag{7.6}$$

We can deduce from Green's formula that

$$\begin{aligned}\left\langle u, \Delta(\chi^2 u)\right\rangle &= \|u\nabla\chi\|^2 - \|\nabla(\chi u)\|^2 \\ &\le \|u\nabla\chi\|^2.\end{aligned}$$

By (7.6) we thus have

$$\|\chi u\|^2 \le \|u\nabla\chi\|^2. \tag{7.7}$$

Now assume that $\chi(x)$ has the form $h(|x| - R)$, where $h(t) = 1$ for $t \le 1$ and $h(t) = 0$ for $t \ge 2$. Letting $R \to \infty$ in (7.7) shows that $u = 0$.

We conclude that $A^* + 1$ is injective. This proves that A is essentially self-adjoint, by the variant of Theorem 3.30 given in Exercise 3.11. □

It is possible to improve the result of Theorem 7.1 by weakening the hypothesis to $V \in L^2_{\mathrm{loc}}(\Omega)$. In this case, the argument that $(-\Delta + V)^* + 1$ is injective relies on a distribution theory result known as Kato's inequality. For details, see, e.g., Hislop and Segal [44, Ch. 8].

7.1.2 Quadratic Form Extension

Another approach to the problem of self-adjoint extension of $-\Delta + V$ is the Friedrichs extension method discussed in Section 3.4.3. Defining extensions in terms of quadratic forms has the great advantage that the form domains are generally simpler than the domains of the corresponding operators.

The Friedrichs method provides a useful complement to the direct arguments used in the previous section. Although Theorem 7.1 gives no information on the exact operator domain, it does establish the crucial property of essential self-adjointness. The Friedrichs method applies more generally and yields a description of the operator domain, but it leaves essential self-adjointness as a separate issue.

Theorem 7.2. *Suppose $V \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is real-valued with $V \ge 0$. Then*

$$\mathcal{H}_Q := \left\{ f \in H^1(\mathbb{R}^n) : \; V^{\frac{1}{2}} f \in L^2(\mathbb{R}^n)\right\} \tag{7.8}$$

is a Hilbert space with respect to the inner product

$$Q[f, g] := \langle f, g\rangle_{H^1} + \langle f, Vg\rangle.$$

There exists a self-adjoint extension of $-\Delta + V$, with domain

$$\mathcal{D}(-\Delta + V) := \Big\{ u \in \mathcal{H}_Q : \; f \mapsto Q[u, f] \textit{ extends to } L^2(\mathbb{R}^n) \textit{ as a bounded functional}\Big\}, \tag{7.9}$$

such that

$$Q[f, g] = \left\langle f, (-\Delta + V + 1)g \right\rangle$$

for all $f, g \in \mathcal{D}(-\Delta + V)$.

Proof Because $V \geq 0$, it is clear that $Q[\cdot, \cdot]$ defines an inner product. To prove completeness, suppose that $\{u_n\} \subset \mathcal{D}(-\Delta + V)$ is a Cauchy sequence with respect to the norm $\|\cdot\|_Q$ associated with Q. Because

$$\|u\|_{H^1} \leq \|u\|_Q,$$

$\{u_n\}$ is Cauchy and thus converges to some $u \in H^1(\mathbb{R}^n)$.

Since

$$\left\|(V + 1)^{\frac{1}{2}} f\right\| \leq \|f\|_Q,$$

$\{(V+1)^{\frac{1}{2}} u_n\}$ is also Cauchy with respect to L^2. Hence there exists some $h \in L^2(\mathbb{R}^n)$ such that

$$(V + 1)^{\frac{1}{2}} u_n \to h \tag{7.10}$$

in the L^2 sense. This implies that $u_n \to (V + 1)^{-\frac{1}{2}} h$ in $L^2(\mathbb{R}^n)$ as well, which implies that $u = (V + 1)^{-\frac{1}{2}} h$. Since h is L^2, we have $u \in \mathcal{H}_Q$, and the combination of H^1 convergence and (7.10) imply that $u_n \to u$ with respect to $\|\cdot\|_Q$. Therefore $\mathcal{H}_Q$ is complete.

For $u \in \mathcal{D}(-\Delta + V)$, the Riesz lemma defines a unique $w \in L^2(\mathbb{R}^n)$ such that

$$Q[u, f] = \langle w, f \rangle$$

for all $f \in \mathcal{H}_Q$. We then define

$$(-\Delta + V)u := w - u.$$

The proof of self-adjointness now follows the same argument used in Theorem 6.6, i.e., we use the Riesz lemma on $\mathcal{H}_Q$ to establish the surjectivity of $-\Delta + V$, which then implies that $\mathcal{D}((-\Delta + V)^*) \subset \mathcal{D}(-\Delta + V)$. □

The space $\mathcal{H}_Q$ defined in (7.8) is called the *form domain* of $-\Delta + V$. Note that although $C_0^\infty(\mathbb{R}^n)$ is clearly contained in $\mathcal{H}_Q$, it is possible that the domain of $-\Delta + V$ given by (7.9) does not include $C_0^\infty(\mathbb{R}^n)$. This is because $V\psi$ need not be contained in $L^2(\mathbb{R}^n)$ for $V \in L^1_{\text{loc}}$ and $\psi \in C_0^\infty(\mathbb{R}^n)$.

Theorem 7.2 illustrates the primary advantage of the Friedrichs method, namely the ease of working with quadratic form domains. As a subspace of $L^2(\mathbb{R}^n)$,

$$\mathcal{H}_Q = H^1(\mathbb{R}^n) \cap \mathcal{D}(M_{V^{1/2}}).$$

In other words, the quadratic form domain of $-\Delta + V$ is just the intersection of the form domains of $-\Delta$ and M_V. On the operator side, the domain $\mathcal{D}(-\Delta + V)$ has no direct connection to $\mathcal{D}(-\Delta) = H^2(\mathbb{R}^n)$ and may in fact be disjoint from it.

When $V \in L^\infty_{\mathrm{loc}}(\mathbb{R}^n)$ (still assuming $V \geq 0$), it is easy to see from (7.9) that $C_0^\infty(\mathbb{R}^n) \subset \mathcal{D}(-\Delta + V)$. Since essential self-adjointness implies uniqueness of the extension, the two methods produce the same result when V satisfies both hypotheses.

7.1.3 Discrete Spectrum

The evolution of states in quantum mechanics is described by the *Schrödinger's equation* for a quantum state Ψ,

$$i \frac{\partial}{\partial t} \Psi = H\Psi, \tag{7.11}$$

where $H = -\Delta + V$ is the quantum Hamiltonian. Eigenfunctions of H correspond to steady-state solutions of the Schrödinger equation. Intuitively, we expect such stable solutions to occur only at energies for which the corresponding classical particle is *trapped*, meaning that its motion is confined to a compact region. Trapping occurs at energies E for which the set $\{V(x) \leq E\}$ is bounded.

In the case that $V \to \infty$ at infinity, classical particles are trapped at all energies. We would therefore expect the Schrödinger operator to have purely discrete spectrum under this condition.

Theorem 7.3. *Suppose $V \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ with $V \geq 0$, and let $-\Delta + V$ be the self-adjoint operator as in Theorem 7.2. If $V(x) \to \infty$ as $x \to \infty$, then $-\Delta + V$ has compact resolvent and purely discrete spectrum.*

Proof Define the Hilbert space $\mathcal{H}_Q$ with quadratic form $Q[\cdot, \cdot]$ as in the proof of Theorem 7.2, and let $A = -\Delta + V$ be the self-adjoint operator defined so that

$$Q[u, v] = \langle u, (A + 1)v \rangle$$

for $u \in \mathcal{D}(A)$ and $v \in \mathcal{H}_Q$. If we set $u = v = (A + 1)^{-1} f$ for $f \in L^2$, then

$$\|u\|_Q^2 = \left\langle (A + 1)^{-1} f, f \right\rangle. \tag{7.12}$$

Suppose $\{f_k\}$ is a bounded sequence in $L^2(\mathbb{R}^n)$. Our goal is to prove that the sequence $u_k := (A + 1)^{-1} f_k$ has a subsequence converging in $L^2(\mathbb{R}^n)$, which will establish that $(A + 1)^{-1}$ is a compact operator.

By (7.12), the sequence u_k is bounded in $\mathcal{H}_Q$. For $m \in \mathbb{N}$, let $\chi_m \in C_0^\infty(\mathbb{R}^n)$ be a cutoff defined so that

$$\chi_m(x) = \begin{cases} 1, & |x| \le m, \\ 0, & |x| \ge m+1. \end{cases}$$

For each m, the sequence $\{\chi_m u_k\}_{k=1}^{\infty}$ is bounded in $H_0^1(B(0; R+1))$, since $\|\cdot\|_{H^1} \le \|\cdot\|_Q$.

Applying Rellich's theorem (Theorem 6.9) to $\{\chi_1 u_k\}$ yields a subsequence $\{u_{1,j}\} \subset \{u_k\}$ for which $\{\chi_1 u_{1,j}\}$ converges in $L^2(\mathbb{R}^n)$. We can then apply Rellich's theorem to $\{\chi_2 u_{1,j}\}$ to produce a subsequence $u_{2,j}$ such that $\{\chi_2 u_{2,j}\}$ converges, and so on. The result is a nested set of subsequences $\{u_{l+1,j}\} \subset \{u_{l,j}\}$ such that $\{\chi_l u_{l,j}\}$ converges in $L^2(\mathbb{R}^n)$ as $j \to \infty$. We then set $w_j := u_{j,j}$. By construction, for each m the cutoff sequence $\{\chi_m w_j\}$ converges in $L^2(\mathbb{R}^n)$ as $j \to \infty$.

Since the sequence $\{u_k\}$ was bounded in $\mathcal{H}_Q$, we also have a uniform bound $\|w_j\|_Q \le M$ for all j. This implies in particular that

$$\left\| V^{\frac{1}{2}} w_j \right\| \le M. \tag{7.13}$$

Given $\varepsilon > 0$, the hypothesis on V allows us to choose m sufficiently large that

$$\inf_{|x| \ge m} V(x) \ge \frac{1}{\varepsilon}.$$

By (7.13) this implies

$$\|(1 - \chi_m) w_j\|^2 \le \varepsilon M^2, \tag{7.14}$$

for all j.

Since $\chi_m w_j$ is convergent in L^2, there exists $N > 0$ such that $i, j \ge N$ implies

$$\|\chi_m (w_i - w_j)\| \le \varepsilon.$$

Combining this with (7.14) gives

$$\begin{aligned} \|w_i - w_j\| &\le \|\chi_m (w_i - w_j)\| + \|(1 - \chi_m) w_i\| + \|(1 - \chi_m) w_j\| \\ &\le \varepsilon + 2M\sqrt{\varepsilon}, \end{aligned}$$

for $i, j \ge N$. Since ε was arbitrary, this demonstrates that $\{w_j\}$ is Cauchy and therefore convergent in $L^2(\mathbb{R}^n)$.

The operator $(A + 1)^{-1}$ is thus compact, and the discreteness of the spectrum follows from Theorem 4.21. □

7.1.4 Quantum Harmonic Oscillator

The classical harmonic oscillator describes the motion of a particle subject to a restoring force proportional to its displacement from the origin. The corresponding potential energy function is quadratic. The quantum mechanical analog of this system (with Planck's constant scaled out) is a Schrödinger operator $-\Delta+V$, acting on $L^2(\mathbb{R}^n)$, with potential given by

$$V(x) = \sum_{i,j=1}^{n} x_i M_{ij} x_j, \tag{7.15}$$

where M is a positive definite matrix. The operator $-\Delta + V$ is essentially self-adjoint on $C_0^\infty(\mathbb{R})$ by Theorem 7.1 and has purely discrete spectrum by Theorem 7.3.

Theorem 7.4. *The spectrum of the quantum harmonic oscillator with potential* (7.15) *is given by*

$$\sigma(-\Delta + V) = \left\{ \sum_{i=1}^{n} (2k_i + 1)\omega_i : k_i \in \mathbb{N}_0 \right\},$$

where $\omega_1^2, \ldots, \omega_n^2$ *denote the eigenvalues of* M *with* $\omega_j > 0$.

By using a linear change of variables to diagonalize M, we can assume that

$$V(x) = \sum_{j=1}^{n} \omega_j^2 x_j^2.$$

Separation of variables then allows us to reduce the eigenvalue equation to a sum of one-dimensional equations. It also suffices to consider the case $\omega_j = 1$, since the coefficient can be accounted for by rescaling the variable. Thus we focus on the one-dimensional operator

$$H := -\Delta + x^2,$$

defined as a self-adjoint operator on $L^2(\mathbb{R})$ by the unique extension described in Theorem 7.1.

To determine the eigenvalues of H, it is convenient to introduce a pair of auxiliary operators,

$$A^\pm := -i\partial_x \pm ix.$$

As differential operators, we have the relations

$$A^+A^- = H - 1, \qquad A^-A^+ = H + 1. \tag{7.16}$$

The operators A^+ and A^- are called *raising* and *lowering* operators, respectively. This terminology is justified by the following:

Lemma 7.5. *Let ψ be an eigenfunction of H with eigenvalue λ. Then ψ is smooth, $A^\pm\psi \in \mathcal{D}(H)$ and*

$$H(A^\pm\psi) = (\lambda \pm 2)A^\pm\psi.$$

Proof First we show that ψ is smooth, by an elliptic regularity argument. For $f \in L^2(\mathbb{R})$, we claim that

$$-\Delta f \in H^m(\mathbb{R}) \text{ implies } f \in H^{m+2}(\mathbb{R}). \tag{7.17}$$

This is an immediate consequence of the characterization of Sobolev spaces (2.26) and the fact that the Fourier transform conjugates $-\Delta$ to multiplication by $|\xi|^2$.

As noted in the proof of Theorem 7.1, $\mathcal{D}(H) \subset H^2_{\text{loc}}(\mathbb{R})$, so this is our starting assumption for ψ. Suppose, for the sake of induction, that $\psi \in H^m_{\text{loc}}(\mathbb{R})$ with $m \geq 2$. For $\chi \in C_0^\infty(\mathbb{R}^n)$, we then have

$$-\Delta(\chi\psi) = (\lambda - x^2)\chi\psi - [\Delta, \chi]\psi.$$

Since $[\Delta, \chi]$ is a first-order differential operator, with smooth coefficients, the right-hand side is contained in $H^{m-1}(\mathbb{R})$ by the inductive hypothesis. Therefore $\chi\psi \in H^{m+1}(\mathbb{R})$, by (7.17). By induction, this shows that $\psi \in H^m_{\text{loc}}(\mathbb{R})$ for all $m \in \mathbb{N}$, which proves that $\psi \in C^\infty(\mathbb{R})$ by Sobolev embedding (Theorem 2.26).

As differential operators, H and $A^\pm$ satisfy a commutator formula,

$$\left[H, A^\pm\right] = \pm 2A^\pm. \tag{7.18}$$

Applying (7.18) to the eigenfunction $\psi \in C^\infty(\mathbb{R}^n)$ gives

$$\begin{aligned} H(A^\pm\psi) &= A^\pm(H \pm 2)\psi \\ &= (\lambda \pm 2)A^\pm\psi. \end{aligned} \tag{7.19}$$

Note that we have only shown that this holds as a differential equation. We still need to check that $A^\pm\psi \in \mathcal{D}(H)$.

By the characterization of the quadratic form domain given in Theorem 7.2, it is clear that $\psi \in H^1(\mathbb{R})$ and $x\psi \in L^2(\mathbb{R})$. Thus we have $A^\pm\psi \in L^2(\mathbb{R})$. For $f \in C_0^\infty(\mathbb{R})$ we can integrate by parts to estimate

$$\begin{aligned} \left|\langle A^\pm\psi, Hf\rangle\right| &= \left|\langle(\lambda \pm 2)a^\pm\psi, f\rangle\right| \\ &\leq C\|f\|. \end{aligned}$$

This shows that $A^\pm\psi$ is contained in $\mathcal{D}(H)$. □

Lemma 7.5 leads rather quickly to an explicit characterization of the spectrum in the dimension one.

Corollary 7.6. *The one-dimensional harmonic oscillator has spectrum*

$$\sigma(-\Delta + x^2) = 2\mathbb{N} - 1,$$

and an orthonormal basis $\{\psi_k\}_{k=0}^{\infty}$ *of eigenfunctions*

$$\psi_k(x) := \pi^{-\frac{1}{4}} 2^{-\frac{k}{2}} (k!)^{-\frac{1}{2}} (\partial_x - x)^k e^{-x^2/2}, \tag{7.20}$$

where $\lambda_k = 2k + 1$.

Proof Suppose that $H\psi = \lambda\psi$ for $\psi \in \mathcal{D}(H)$. Since ψ is smooth, we can integrate by parts to obtain

$$\begin{aligned} \|A^-\psi\|^2 &= \langle \psi, A^+A^-\psi \rangle \\ &= \langle \psi, (H - I)\psi \rangle \\ &= (\lambda - 1)\|\psi\|^2. \end{aligned} \tag{7.21}$$

Iterating this result gives

$$\left\|(A^-)^k \psi\right\|^2 = \|\psi\|^2 \prod_{j=1}^{k} (\lambda + 1 - 2j) \tag{7.22}$$

for $k \in \mathbb{N}$. If $\lambda \notin 2\mathbb{N} - 1$, then the product on the right is strictly negative for large values of k, which is impossible for $\psi \neq 0$. Therefore $\sigma(-\Delta + x^2)$ is contained in the set $2\mathbb{N} - 1$.

If ψ_0 satisfies the eigenvalue equation for $\lambda_0 = 1$, then $A^-\psi_0 = 0$ by (7.21). Thus ψ_0 satisfies

$$\frac{d}{dx}\psi_0 = x\psi_0,$$

which has the L^2-normalized solution,

$$\psi_0(x) := \pi^{-\frac{1}{4}} e^{-x^2/2},$$

unique up to a multiplicative constant. The uniqueness shows that $\lambda_0 = 1$ is a simple eigenvalue.

By Lemma 7.5, $(A^+)^k\psi_0$ is an eigenfunction for $\lambda_k := 2k + 1$, and its normalization can be calculated using (7.16),

$$\left\| (A^+)^k \psi_0 \right\|^2 = \prod_{l=0}^{k-1} (\lambda_l + 1)$$
$$= 2^k k!$$

Normalizing $(A^+)^k \psi_0$ thus yields the eigenfunction ψ_k given by (7.20).

To see that the higher eigenvalues are also simple, note that if ψ is an eigenfunction with eigenvalue λ_k, then $(A^-)^k \psi$ has eigenvalue 1. Since λ_0 was simple, this implies that $(A^-)^k \psi$ is a constant multiple of ψ_0, and it follows that ψ is proportional to ψ_k. □

A few eigenfunctions of H are illustrated in Figure 7.1. Each eigenfunction consists of a polynomial factor times $e^{-x^2/2}$. These factors are (up to a normalization constant) called the *Hermite polynomials*.

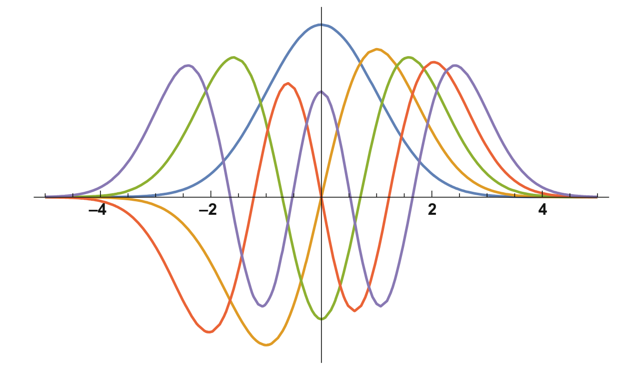


Fig. 7.1 The first five eigenfunctions of the one-dimensional harmonic oscillator

The spectrum of $-\Delta + \omega^2 x^2$ on $L^2(\mathbb{R})$ is derived from Corollary 7.6 by a scaling argument. Consider the unitary transformation of $L^2(\mathbb{R})$ defined by

$$U_\omega f(x) := \omega^{\frac{1}{4}} f(\gamma^{\frac{1}{2}} x).$$

Conjugating the operator H by U_ω gives

$$U_\omega(-\Delta + x^2)U_\omega^{-1} = -\omega^{-1}\Delta + \omega x^2.$$

So that

$$-\Delta + \omega^2 x^2 = \omega U_\omega(-\Delta + x^2)U_\omega^{-1}.$$

Hence, by Corollary 7.6,

$$\sigma\left(-\Delta + \omega^2 x^2\right) = (2\mathbb{N} - 1)\omega. \tag{7.23}$$

Theorem 7.4 now follows from (7.23) by diagonalization and separation of variables.

7.2 Relatively Bounded Perturbations

The methods of Section 7.1 are not applicable to potentials describing electromagnetic interactions between particles, because the Coulomb potential (7.2) is not semi-bounded below. In this section we will develop a general method that establishes self-adjointness in cases like the Coulomb potential, by regarding the potential term as a relatively small perturbation of $-\Delta$.

To make this notion precise, we say that a symmetric operator B is *relatively bounded* with respect to a self-adjoint operator A if $\mathcal{D}(A) \subset \mathcal{D}(B)$ and there exist constants $\alpha, \beta \in \mathbb{R}$ such that

$$\|Bu\| \leq \alpha\|Au\| + \beta\|u\| \tag{7.24}$$

for all $u \in \mathcal{D}(A)$. The following result was proven by Franz Rellich in 1939 [73], and its applications to Schrödinger operators in particular were developed by Tosio Kato [50].

Theorem 7.7 (Kato–Rellich). *Let A be a self-adjoint operator on a Hilbert space $\mathcal{H}$. If B is a symmetric operator which is relatively bounded with respect to A, with constant $\alpha < 1$ in (7.24), then $A + B$ is self-adjoint on $\mathcal{D}(A)$ and essentially self-adjoint on any core for A.*

Proof Our goal is to apply Theorem 3.29 by establishing the surjectivity of $A + B - i\sigma$ for $\sigma \in \mathbb{R}$ with $|\sigma|$ large. Since A is self-adjoint, $A - i\sigma$ is invertible for $\sigma \neq 0$. By the assumption that $\mathcal{D}(A) \subset \mathcal{D}(B)$ we can thus write

$$A + B - i\sigma = \left[I + B(A - i\sigma)^{-1}\right](A - i\sigma), \tag{7.25}$$

and argue for surjectivity by estimating $B(A - i\sigma)^{-1}$.

For $u \in \mathcal{D}(A)$, the assumption (7.24) gives

$$\left\|B(A - i\sigma)^{-1}u\right\| \leq \alpha\left\|A(A - i\sigma)^{-1}u\right\| + \beta\|(A - i\sigma)^{-1}u\|. \tag{7.26}$$

The spectral theorem gives the estimates

$$\left\|A(A - i\sigma)^{-1}\right\| \leq 1, \qquad \left\|(A - i\sigma)^{-1}\right\| \leq |\sigma|^{-1}.$$

Hence, (7.26) implies that $B(A - i\sigma)^{-1}$ is a bounded operator with

$$\left\|B(A - i\sigma)^{-1}\right\| \leq \alpha + \beta|\sigma|^{-1}.$$

By hypothesis $\alpha < 1$, so by taking $|\sigma|$ sufficiently large, we can assume that

$$\left\|B(A - i\sigma)^{-1}\right\| < 1.$$

This guarantees that $I + B(A - i\sigma)^{-1}$ is invertible by Neumann series. It then follows from (7.25) that $A + B \pm i\sigma$ is surjective as a map $\mathcal{D}(A) \to \mathcal{H}$, provided $|\sigma|$ is large enough. Therefore, by Theorem 3.29, $A + B$ is self-adjoint on $\mathcal{D}(A)$.

Now assume that A is merely essentially self-adjoint. If $u \in \mathcal{D}(\overline{A})$, then there exists a sequence $u_n \to u$ such that Au_n converges to $\overline{A}u$. By the assumption (7.24), the sequence Bu_n also converges, so that $u \in \mathcal{D}(B)$ (since B is closed). By continuity, we can extend (7.24) to

$$\left\|\overline{B}u\right\| \le \alpha \left\|\overline{A}u\right\| + \beta\|u\| \tag{7.27}$$

for all $u \in \mathcal{D}(\overline{A})$. By the first part of the proof, this implies that $\overline{A} + B$ is self-adjoint on the domain $\mathcal{D}(\overline{A})$.

Using the fact that $(A + B)u_n \to \overline{A}u + Bu$, we can also conclude that $u \in \mathcal{D}(\overline{(A + B)})$, with

$$\overline{(A + B)}u = \overline{A}u + Bu.$$

This means that

$$\overline{(A + B)} \subset \overline{A} + B.$$

Because $\overline{A} + B$ is self-adjoint, it is closed in particular. Thus, since $\overline{A} + B$ a closed extension of $A + B$ and $\overline{(A + B)}$ is the smallest closed extension, we have

$$\overline{A} + B \subset \overline{(A + B)}.$$

We conclude that $\overline{(A + B)} = \overline{A} + B$, which is self-adjoint. Thus $A + B$ is essentially self-adjoint. □

Kato's original application of Theorem 7.7 to Schrödinger operators includes the following result, which covers the case of the Coulomb potential in particular.

Theorem 7.8. *Let $n \le 3$, and suppose V is a real-valued potential in $L^2(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$. Then $-\Delta + V$ is essentially self-adjoint on $C_0^\infty(\mathbb{R}^n)$ and self-adjoint on $H^2(\mathbb{R}^n)$.*

Proof The $L^\infty(\mathbb{R}^n)$ component is included for the sake of applications, but makes no difference to the proof because it contributes a bounded self-adjoint operator. Therefore, it suffices to consider $V \in L^2(\mathbb{R}^n)$.

Suppose that $u \in H^2(\mathbb{R}^n)$. By the Sobolev embedding (Theorem 2.26), this implies that u is continuous and bounded for $n < 4$. This gives $Vu \in L^2(\mathbb{R}^n)$. Since this is the defining condition for the domain of the multiplication operator M_V, we conclude that

$$H^2(\mathbb{R}^n) \subset \mathcal{D}(M_V).$$

As in the proof of Sobolev embedding, we can use the fact that

$$\|u\|_\infty \le (2\pi)^{-n/2}\|\hat{u}\|_1,$$

together with the Cauchy–Schwarz inequality, to estimate

$$\|u\|_\infty^2 \le C\int_{\mathbb{R}^n}\frac{1}{(|\xi|^2+b^2)^2}d^n\xi \cdot \int_{\mathbb{R}^n}(|\xi|^2+b^2)^2|\hat{u}(\xi)|^2 d^n\xi, \tag{7.28}$$

for $b > 0$. The first integral gives

$$\int_{\mathbb{R}^n}\frac{1}{(|\xi|^2+b^2)^2}d^n\xi = Cb^{n-4},$$

where C depends only on n, while the second is equal to $\|(-\Delta+b^2)u\|^2$. Therefore, (7.28) gives the estimate

$$\|u\|_\infty \le Cb^{n-4}\big\|(-\Delta+b^2)u\big\|. \tag{7.29}$$

For $V \in L^2(\mathbb{R}^n)$ and $u \in H^2(\mathbb{R}^n)$, we apply (7.29) to obtain

$$\begin{aligned}\|Vu\| &\le \|V\|\|u\|_\infty \\ &= Cb^{n-4}\|V\|\Big(\|\Delta u\| + b^2\|u\|\Big),\end{aligned}$$

for all $b > 0$. Taking b sufficiently large gives

$$\|Vu\| \le \alpha\|\Delta u\| + \beta\|u\|,$$

with α arbitrarily small. Thus $-\Delta + V$ satisfies the hypotheses of Theorem 7.7. □

Example 7.9. On $\mathbb{R}^3$, consider the Hamiltonian operator for the hydrogen atom $H := -\Delta - r^{-1}$. The potential $V(x) = -r^{-1}$ lies in $L^2_{\text{loc}}(\mathbb{R}^3)$ and is bounded outside a compact set. Therefore, H is essentially self-adjoint on $C_0^\infty(\mathbb{R}^3)$ and self-adjoint on $H^2(\mathbb{R}^3)$ by Theorem 7.8. ♢

Example 7.10. For a charged particle in a magnetic field, the quantum Hamiltonian associated with the magnetic vector potential $A : \mathbb{R}^3 \to \mathbb{R}^3$ is

$$H = (i\nabla + A)^2, \tag{7.30}$$

where the physical constants are omitted. An operator of this type is called a *magnetic Schrödinger operator*.

For $f, g \in C_0^\infty(\mathbb{R}^3)$, we have

$$\langle f, Hg\rangle = \langle (i\nabla + A)f, (i\nabla + A)g\rangle,$$

which shows that H is symmetric as an operator on $C_0^\infty(\mathbb{R}^3)$. Expanding the square gives

$$H = -\Delta + 2iA \cdot \nabla + (\nabla \cdot A) + A^2.$$

The scalar terms $(\nabla \cdot A)$ and A^2 are bounded, provided we assume that A and $\nabla \cdot A$ lie in $L^\infty(\mathbb{R}^3)$, (where the divergence is defined in the weak sense). However, the gradient term $2iA \cdot \nabla$ is an unbounded operator.

To apply Theorem 7.7, note that for $f \in H^2(\mathbb{R}^3)$,

$$\begin{aligned}\|\nabla f\|^2 &= \|\xi \hat{f}\|^2 \\ &= \int_{\mathbb{R}^n} |\xi|^2 |\hat{f}(\xi)|^2 \, d^n\xi \\ &\le \frac{1}{2} \int_{\mathbb{R}^n} (1 + |\xi|^4) |\hat{f}(\xi)|^2 \, d^n\xi \\ &= \frac{1}{2}\Big(\|f\|^2 + \|\Delta f\|^2\Big).\end{aligned}$$

For $A \in L^\infty(\mathbb{R}^3)$, it follows that $2iA \cdot \nabla$ is relatively bounded with respect to $-\Delta$. Therefore, H is self-adjoint on $H^2(\mathbb{R}^3)$ and essentially self-adjoint on $C_0^\infty(\mathbb{R}^3)$. ◊

7.3 Relatively Compact Perturbations

In this section we take up the question of locating the essential spectrum of $-\Delta + V$. Unfortunately, Weyl's stability result (Theorem 5.14), which says the essential spectrum is unaffected by compact perturbations, is not helpful for Schrödinger operators. The multiplication operator M_V is not compact on $L^2(\mathbb{R}^n)$ unless V vanishes almost everywhere.

Our first goal is therefore to strengthen the Weyl stability result. Instead of requiring the difference in operators to be compact, we can impose this condition on the difference of resolvents.

Theorem 7.11. *Suppose A and B are self-adjoint operators on a Hilbert space $\mathcal{H}$. If there exists $z \in \rho(A) \cap \rho(B)$ such that $(A - z)^{-1} - (B - z)^{-1}$ is compact, then $\sigma_{\mathrm{ess}}(A) = \sigma_{\mathrm{ess}}(B)$.*

Proof Suppose that $\lambda \in \sigma_{\mathrm{ess}}(A)$. By Theorem 5.13, there exists a sequence $\{u_k\} \subset \mathcal{D}(A)$ with $\|u_k\| = 1$ and $u_k \to 0$ in the weak sense, such that

$$\lim_{k\to\infty} (A - \lambda)u_k = 0. \tag{7.31}$$

By the definition of the resolvent,

$$(A - z)^{-1}(A - \lambda) = I - (z - \lambda)(A - z)^{-1},$$

so (7.31) implies that

$$\lim_{k\to\infty} \left\| u_k - (z - \lambda)(A - z)^{-1} u_k \right\| = 0. \tag{7.32}$$

The assumption that $(A - z)^{-1} - (B - z)^{-1}$ is compact implies that

$$\lim_{k\to\infty} \left[(A - z)^{-1} - (B - z)^{-1} \right] u_k = 0,$$

by Theorem 3.38. Therefore, we can deduce from (7.32) that

$$\lim_{k\to\infty} \left\| u_k - (z - \lambda)(B - z)^{-1} u_k \right\| = 0. \tag{7.33}$$

If we now set

$$w_k := (B - z)^{-1} u_k,$$

then (7.33) translates to

$$\lim_{k\to\infty} \|(B - \lambda) w_k\| = 0. \tag{7.34}$$

By writing (7.33) in the form

$$\lim_{k\to\infty} \|u_k - (z - \lambda) w_k\|,$$

and using the fact that $\|u_k\| = 1$, we can also deduce that

$$\lim_{k\to\infty} \|w_k\| = |z - \lambda|^{-1}, \tag{7.35}$$

by (7.33).

Suppose that $\lambda \neq \sigma_{\mathrm{ess}}(B)$. If Π denotes the spectral resolution of B, then this means that $\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}$ has finite rank for some $\varepsilon > 0$. In particular, $\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}(B - z)^{-1}$ will be compact, so that

$$\lim_{k\to\infty} \Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)} w_k = 0. \tag{7.36}$$

With $Q_\varepsilon := I - \Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}$, we can estimate

$$\begin{aligned}\|(B-\lambda)w_k\| &= \left\|(B-\lambda)(\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}w_k + Q_\varepsilon w_k)\right\| \\ &\geq \|(B-\lambda)Q_\varepsilon w_k\| - \left\|(B-\lambda)\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}w_k\right\| \\ &\geq \varepsilon\|Q_\varepsilon w_k\| - \varepsilon\left\|\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}w_k\right\|.\end{aligned}$$

Taking $k \to \infty$, using (7.36) and (7.35), gives

$$\liminf_{k\to\infty}\|(B-\lambda)w_k\| \geq \varepsilon|z-\lambda|.$$

This contradicts (7.34), so we conclude that $\lambda \in \sigma_{\text{ess}}(B)$.

We have shown that $\sigma_{\text{ess}}(A) \subset \sigma_{\text{ess}}(B)$, and the same argument applies with A and B switched. □

In many applications of Theorem 7.11, including Schrödinger operators, only one of the resolvents is explicitly known. It is thus helpful to invoke the second resolvent identity (Exercise 4.1), which says that

$$(B-z)^{-1} - (A-z)^{-1} = (B-z)^{-1}(B-A)(A-z)^{-1}. \tag{7.37}$$

To satisfy the hypothesis of Theorem 7.11, we can see from (7.37) that it is sufficient to verify the compactness of $(B-A)(A-z)^{-1}$.

With this in mind, we say that an operator T is *relatively compact* with respect to A if $\mathcal{D}(A) \subset \mathcal{D}(T)$ and $T(A-z)^{-1}$ is compact for some $z \in \rho(A)$. By the first resolvent identity (Corollary 4.12), if $T(A-z)^{-1}$ is compact for one value of z, then it is compact for all $z \in \rho(A)$.

Theorem 7.12 (Improved Weyl Stability). *If A and T are self-adjoint operators and T is relatively compact with respect to A, then $A+T$ is self-adjoint on $\mathcal{D}(A)$ and*

$$\sigma_{\text{ess}}(A+T) = \sigma_{\text{ess}}(A).$$

Proof First consider the self-adjointness of $A+T$. Our plan is to show that $\|T(A+i\sigma)^{-1}\|$ is small for large σ, and then argue as in the proof of Kato–Rellich (Theorem 7.7).

Through the functional calculus, the operator $(A+i)(A-i\sigma)^{-1}$ corresponds to multiplication by the function

$$f_\sigma(x) := \frac{x+i}{x-i\sigma},$$

for $x \in \mathbb{R}$. Note that $|f_\sigma(x)| \leq 1$ and $f_\sigma \to 0$ pointwise as $\sigma \to \infty$. Therefore, using the dominated convergence as in Example 2.12, we have

$$(A+i)(A-i\sigma)^{-1} \to 0 \tag{7.38}$$

in the strong operator sense as $\sigma \to 0$.

We can now write

$$T(A+i\sigma)^{-1} = T(A-i)^{-1}(A-i)(A+i\sigma)^{-1}.$$

Since $T(A-i)^{-1}$ is compact by assumption, and (7.38) gives strong convergence

$$\left[(A-i)(A+i\sigma)^{-1}\right]^* \to 0,$$

it follows that

$$\lim_{\sigma\to 0} \left\| T(A+i\sigma)^{-1} \right\| = 0, \tag{7.39}$$

by the result from Exercise 3.14. By choosing $|\sigma|$ sufficiently large, we can ensure that

$$\left\| T(A+i\sigma)^{-1} \right\| < 1.$$

As in the proof of Theorem 7.7, this implies that $A+T\pm i\sigma$ is surjective, and hence that $A+T$ is self-adjoint on $\mathcal{D}(A)$.

The fact that A and $A+T$ have the same essential spectra now follows immediately from Theorem 5.14, by (7.37). □

Suppose $-\Delta + V$ is self-adjoint on a domain that includes $H^1(\mathbb{R}^n)$, as is the case for the potentials considered in Section 7.1.2, for example. Then, for $\chi \in C_0^\infty(\mathbb{R}^n)$, Rellich's theorem (Theorem 6.9) implies that $\chi(-\Delta+V)^{-1}$ is compact on $L^2(\mathbb{R}^n)$. This observation yields the following:

Corollary 7.13. *Suppose that V is a potential such that $-\Delta + V$ is self-adjoint on a domain that includes $H^1(\mathbb{R}^n)$. If $W \in L^\infty(\mathbb{R}^n; \mathbb{R})$ and has compact support, then*

$$\sigma_{\mathrm{ess}}(-\Delta + V) = \sigma_{\mathrm{ess}}(-\Delta + V + W).$$

Example 7.14. On $\mathbb{R}$, consider the Schrödinger operator $H := -\Delta + V$ with the negative square-well potential

$$V(x) = \begin{cases} -c, & |x| \le a, \\ 0, & |x| > a, \end{cases}$$

where a, c are positive constants. By Theorem 7.8, H is self-adjoint with domain $H^2(\mathbb{R})$. Since V has compact support, Corollary 7.13 shows that the essential spectrum is $[0,\infty)$.

To find the discrete spectrum, note that for $\lambda \ge 0$ there are no L^2 solutions of the eigenvalue equation for $|x| > a$, so all eigenvalues are negative. Clearly $-\Delta + c \ge 0$, so the spectrum is bounded below by $-c$. Since the potential is symmetric, we can assume, by averaging if needed, that all eigenfunctions have either even or odd symmetry.

First consider the even eigenfunctions for $\lambda = -\sigma^2$ with $0 < \sigma < \sqrt{c}$. For $|x| > a$ the L^2 requirement implies the eigenfunctions take the form $\psi(x) = e^{-\sigma|x|}$. On $[-a, a]$, an even solution is a constant multiple of $\cos(\omega x)$, where $\omega := \sqrt{|c - \sigma^2|}$. Since functions in $H^2(\mathbb{R})$ are C^1 in particular, the first derivatives will match at $x = \pm a$. The matching conditions at $x = \pm a$ reduce to

$$\sigma = \omega \tan(\omega a).$$

Because $\omega \tan(\omega a)$ is decreasing as a function of σ, and equal to zero at $\sigma = \sqrt{c}$, there is at least one intersection point for $0 < \sigma < \sqrt{c}$. Hence an even eigenfunction exists for all $c > 0$, and H has at least one negative eigenvalue. There are multiple even eigenfunctions if $c > (\pi/a)^2$.

For odd eigenfunctions, the matching condition is

$$\sigma = -\omega \cot(\omega a).$$

An odd eigenfunction therefore exists if and only if $c > (\pi/2a)^2$.

The eigenfunctions concentrate on the support of V, as illustrated in Figure 7.2. See Exercise 7.6 for a more general exploration of this phenomenon. ♢

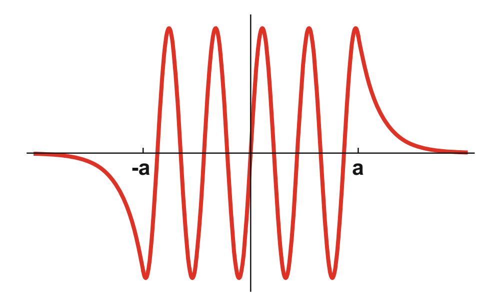


Fig. 7.2 Sample eigenfunction for the square-well potential supported on $[-a, a]$

Theorem 7.12 can be applied to more general potentials in cases where we have better knowledge of the resolvent. As an example, let us consider the class of potentials covered by Theorem 7.8, which were assumed to be $L^2 + L^\infty$, with the dimension less than or equal to three. The first step is to use our explicit knowledge of the Green's function to establish relative compactness for L^2 potentials.

Lemma 7.15. *If $V \in L^2(\mathbb{R}^n)$ with $n \leq 3$, then M_V is relatively compact with respect to $-\Delta$, and therefore*

$$\sigma_{\mathrm{ess}}(-\Delta + V) = [0, \infty).$$

Proof The computation of the resolvent for $\mathbb{R}^n$ in Section 4.1.3 gives the integral kernel $(-\Delta + 1)^{-1}$ as

$$G_n(r) := (2\pi)^{-\frac{n}{2}} r^{1-\frac{n}{2}} K_{\frac{n}{2}-1}(r),$$

where $r = |x|$. From standard Bessel asymptotics [64, §10.30], we have

$$K_\nu(r) = \begin{cases} O(r^{-|\nu|}), & \nu \neq 0, \\ O(\log r), & \nu = 0, \end{cases}$$

as $r \to 0$, and $K_\nu(r) = O(e^{-r})$ as $r \to \infty$. This gives

$$G_n(r) = \begin{cases} O(r), & n = 1, \\ O(\log r), & n = 2, \\ O(r^{2-n}), & n \geq 3. \end{cases}$$

In combination with the exponential decay at infinity, these estimates imply that

$$G_n \in L^2(\mathbb{R}^n), \tag{7.40}$$

provided $n \leq 3$.

The integral kernel of $V(-\Delta + 1)^{-1}$ is

$$K(x, y) = V(x)G_n(|x - y|).$$

By (7.40), we can estimate

$$\begin{aligned} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |K(x, y)|^2 d^n x \, d^n y &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \left| V(x)^2 G_n(|x - y|) \right|^2 d^n x \, d^n y \\ &= \|V\|^2 \, \|G_n\|^2 \\ &< \infty. \end{aligned}$$

This shows that $V(-\Delta + 1)^{-1}$ is a Hilbert–Schmidt operator, which is compact by Theorem 3.39.

We have proven relative compactness, and the characterization of the essential spectrum follows from Theorem 7.12. □

In order to extend the result of Lemma 7.15 to potentials in the class $L^2 + L^\infty$, including the Coulomb potential, we need to add an extra condition. Shifting the operator by a constant will obviously move the essential spectrum. Thus, in order to control the result we must limit the bounded contribution to the potential by imposing decay at infinity. The rate at which the potential decays turns out to be irrelevant; the limit $V \to 0$ is sufficient.

Theorem 7.16. *Let V be a real-valued potential in $L^2_{\text{loc}}(\mathbb{R}^n)$ for $n \leq 3$. If $V(x) \to 0$ as $|x| \to \infty$, then M_V is relatively compact with respect to $-\Delta$ and therefore*

$$\sigma_{\text{ess}}(-\Delta + V) = [0, \infty).$$

Proof Let χ_n denote the characteristic function of the set $\{|x| \leq n\}$ for $n \in \mathbb{N}$. Since

$$\left\|(-\Delta + 1)^{-1}\right\| = 1,$$

by the resolvent estimate (5.20), we have an operator norm bound

$$\left\|(1 - \chi_n)V(-\Delta + 1)^{-1}\right\| \leq \sup_{|x|\geq n} V(x).$$

By the assumption that $V \to 0$, this implies that

$$\chi_n V(-\Delta + 1)^{-1} \to V(-\Delta + 1)^{-1}$$

as $n \to \infty$, with convergence in operator norm. Since $\chi_n V(-\Delta + 1)^{-1}$ is compact, by Lemma 7.15, it follows that $V(-\Delta + 1)^{-1}$ is compact by Theorem 3.36. The claimed result now follows from Theorem 7.12. □

Theorem 7.16 applies in particular to the quantum Hamiltonian of the hydrogen atom, introduced in Example 7.9. We will work out the spectrum of this operator explicitly in the next section. There are results analogous to Theorem 7.16 in higher dimensions, but the proofs are more difficult. In dimension greater than three we do not have the convenience of exploiting the Hilbert–Schmidt property to prove compactness.

7.4 Hydrogen Atom

We have now all the tools in places to analyze the spectrum of the quantum Hamiltonian for the hydrogen atom,

$$H = -\Delta - \frac{1}{r}, \tag{7.41}$$

acting on $L^2(\mathbb{R}^3)$. By Theorem 7.8, H is essentially self-adjoint on $C_0^\infty(\mathbb{R}^3)$ and self-adjoint on $H^2(\mathbb{R}^3)$. Furthermore, Theorem 7.16 shows that

$$\sigma_{\text{ess}}(H) = [0, \infty). \tag{7.42}$$

This continuous range of positive energies corresponds to ionized states of the atom, where the electron has broken free from the nucleus.

In this section, we will show that the discrete spectrum of H consists of a sequence of negative eigenvalues accumulating at zero. These are the "bound states" of the atom, for which the electron is effectively trapped within some neighborhood

of the nucleus. The absorption and emission lines of atomic hydrogen correspond to the difference in energy between two bound states. An atom can only absorb or emit photons at these particular frequencies.

Theorem 7.17. *The hydrogen operator* (7.41) *has discrete spectrum consisting of negative eigenvalues,*

$$\lambda_n = -\frac{1}{4n^2},$$

for $n \in \mathbb{N}$, each with multiplicity n^2.

Proof By (7.42), the discrete spectrum is negative. Suppose that $\lambda < 0$ is an eigenvalue, and let $W_\lambda \subset H^2(\mathbb{R}^3)$ be the corresponding eigenspace, which is finite-dimensional.

In spherical coordinates (r, φ, θ), the Laplacian takes the form

$$\Delta = \frac{1}{r^2}\partial_r\Big(r^2\partial_r\Big) + \frac{1}{r^2\sin\varphi}\partial_\varphi\big(\sin\varphi\,\partial_\varphi\big) + \frac{1}{r^2\sin^2\varphi}\partial_\theta^2. \tag{7.43}$$

Note that the angular variables can be separated out by defining a *spherical Laplacian*

$$\Delta_{\mathbb{S}^2} := \frac{1}{\sin\varphi}\partial_\varphi\big(\sin\varphi\,\partial_\varphi\big) + \frac{1}{\sin^2\varphi}\partial_\theta^2. \tag{7.44}$$

Here $\mathbb{S}^2$ denotes the unit sphere in $\mathbb{R}^3$. From the relation,

$$\Delta = \frac{1}{r^2}\partial_r\Big(r^2\partial_r\Big) + \frac{1}{r^2}\Delta_{\mathbb{S}^2},$$

it is clear that Δ and $\Delta_{\mathbb{S}^2}$ commute as differential operators.

The action of $\Delta_{\mathbb{S}^2}$ on W_λ is well defined. Indeed, a local regularity analysis as in the proof of Lemma 7.5 can be used to show that eigenfunctions of H are smooth except possibly at the origin. Since $\Delta_{\mathbb{S}^2}$ commutes with Δ, it maps W_λ to itself. Moreover, using integration by parts in spherical coordinates, we can easily check that $-\Delta_{\mathbb{S}^2}$ is a positive operator on the domain $H^2(\mathbb{R}^3)$. Therefore the restriction of $\Delta_{\mathbb{S}^2}$ to W_λ is self-adjoint, since this space is finite-dimensional.

The same reasoning applies to the operator $-i\partial_\theta$, which commutes with both Δ and $\Delta_{\mathbb{S}^2}$ and restricts to a self-adjoint operator on W_λ. (A physicist would interpret $-\Delta_{\mathbb{S}^2}$ and $-i\partial_\theta$ as quantum observables in terms of angular momentum. The square of the total angular momentum is represented by $-\Delta_{\mathbb{S}^2}$, while $-i\partial_\theta$ gives the component of angular momentum about the z-axis.)

By the spectral theorem in its finite-dimensional form, W_λ admits a basis consisting of simultaneous eigenfunctions of $-\partial_\theta$ and $\Delta_{\mathbb{S}^2}$. Suppose that u is such

a joint eigenfunction. Since u is required to be smooth away from the origin, it is clear from the fact that u is a periodic eigenfunction of $-i\partial_\theta$ that u has the form

$$u(r, \varphi, \theta) = e^{im\theta} f(r, \varphi),$$

for $m \in \mathbb{Z}$ and some function f on $[0, \infty) \times [0, \pi]$. Suppose now that

$$-\Delta_{\mathbb{S}^2} u = \kappa u$$

for $\kappa > 0$. After making the substitutions $u = e^{im\theta} f$ and $z = \cos\varphi$, this eigenvalue equation becomes

$$(1 - z^2) f'' - 2z f' + \left(\kappa - \frac{m^2}{1 - z^2}\right) f = 0.$$

This is the Legendre equation; a pair of linearly independent solutions is given by the *Legendre functions* $P_l^m(z)$ and $Q_l^m(z)$, where $\kappa = l(l+1)$.

For an element of W_λ we require $e^{im\theta} f(r, \varphi)$ to be smooth away from the origin. Since the θ coordinate is singular on the z-axis, represented by $z = \pm 1$, we need the Legendre solutions to be regular at these points. The solution $Q_l^m(z)$ diverges as $z \to 1$ for all choices of parameter, so this is ruled out. The solution $P_l^m(z)$ is regular at $z = 1$, but diverges as $z \to -1$, except in the special case where $l \in \mathbb{N}_0$ with $|m| \leq l$. Assuming that l, m satisfy these conditions, the Rodrigues formula [64, §14.7] gives

$$P_l^m(z) = \frac{(-1)^m}{2^l l!} (1 - z^2)^{m/2} \frac{d^{l+m}}{dz^{l+m}} (z^2 - 1)^l. \tag{7.45}$$

From this expression it is easy to check that $P_l^m(\cos\varphi)$ is a polynomial in $\sin\varphi$ and $\cos\varphi$.

With these considerations, we have narrowed the choices for u down to the form

$$u(r, \varphi, \theta) = e^{im\theta} P_l^m(\cos\varphi) h(r), \tag{7.46}$$

for $l \in \mathbb{N}_0$, $m \in \{-m, \ldots, m\}$, and h a function on $[0, \infty)$. The combination $e^{im\theta} P_l^m(\cos\varphi)$ is called a *spherical harmonic*, and traditionally denoted by $Y_l^m(\varphi, \theta)$, up to a normalization constant.

We turn now to the radial equation. Since (7.46) implies that $-\Delta_{\mathbb{S}^2} u = l(l+1)u$, the full eigenvalue equation $Hu = \lambda u$ reduces to the radial form

$$\left[-\frac{1}{r^2}\partial_r\left(r^2\partial_r\right) - \frac{1}{r} + \frac{l(l+1)}{r^2}\right] h = \lambda h. \tag{7.47}$$

Multiplication by r^2 yields a second order ODE with polynomial coefficients,

$$h'' + 2rh' + \Big(r + \lambda r^2 - l(l+1)\Big)h = 0,$$

to which the Frobenius method of power series can be applied. The indicial roots are l and $-l-1$. Since these differ by an integer, the theorem of Fuchs (see, e.g., Teschl [91, Thm. 4.5]) gives two linearly independent solutions of the form,

$$h_1(r) = r^l f_1(r), \qquad h_2(r) = r^{-l-1} f_2(r) + cr^l f_1(r)\log r,$$

where the functions f_j are entire and satisfy $f_j(0) = 1$, and c is a constant. The singular solution h_2 can be ruled out in our case, because this would yield a function u which does not lie in $H^2(\mathbb{R}^3)$. Therefore, the radial component of an eigenfunction has a power series expansion with infinite radius of convergence, starting from the term r^l.

To simplify the calculation of coefficients, it is useful to extract the asymptotic behavior as $r \to 0$. The leading part of (7.47) at infinity is $-h'' = \lambda h$. If we set $\lambda = -\sigma^2$ with $\sigma > 0$, then the expected asymptotic behavior is $e^{\pm\sigma r}$. Of course, only the decaying solution would be compatible with an H^2 function. Therefore, we take the ansatz

$$h(r) = q(r)r^l e^{-\sigma r}, \tag{7.48}$$

and seek a power series expansion of the form

$$q(r) = \sum_{k=0}^{\infty} a_k r^k.$$

Plugging (7.48) into (7.47) and identifying the coefficients of r^{k-1} yield a recursive relation

$$a_k = \frac{2\sigma(k+l) - 1}{k(k+2l+1)} a_{k-1}. \tag{7.49}$$

The resulting series is convergent for any value of σ. However, for $\delta > 0$ the ratio a_k/a_{k-1} is bounded below by $(2-\delta)\sigma/k$ for k sufficiently large. This implies an exponential lower bound $q(r) \geq ce^{(2-\delta)r}$, unless the sequence defined by (7.49) terminates with some coefficient equal to zero. A lack of termination would imply that $h(r)$ grows exponentially as $r \to \infty$, which is obviously ruled out by the L^2 requirement.

Therefore, for $h(r)$ to be the radial component of an eigenfunction, we must have $\sigma = 1/2n$ for some integer $n > l$, so that all terms in the sequence determined by (7.49) vanish for $k \geq n - l$. If the resulting polynomial is denoted by $q_{n,l}(r)$, then our potential eigenfunction has the form

$$u_{n,l,m}(r, \varphi, \theta) = r^l q(r) Y_l^m(\varphi, \theta) e^{-r/2n},$$

Clearly $u_{n,l,m}$ decays exponentially as $r \to \infty$. Moreover, $r^l Y_l^m(\varphi, \theta)$ is a homogeneous polynomial of degree l in the variables x, y, z, so it is easy to check that $u_{n,l,m} \in H^2(\mathbb{R}^3)$.

The eigenvalue corresponding to $u_{n,l,m}$ is $\lambda = -\sigma^2$, so the discrete spectrum of H is given by

$$\sigma_{\text{disc}}(H) = \left\{ -\frac{1}{4n^2} : n \in \mathbb{N} \right\}.$$

For each value of n, the possible choices are $l \in \{0, \ldots, n-1\}$ and $m \in \{-l, \ldots, l\}$. The multiplicity of $-1/4n^2$ is thus

$$\sum_{l=0}^{n-1} (2l+1) = n^2.$$

□

If we reintroduce the physical constants, as in (7.1), the quantum Hamiltonian becomes

$$H = -\frac{\hbar^2}{2m}\Delta - \frac{kq^2}{r}.$$

By rescaling the eigenvalues from Theorem 7.17, we can compute the energy levels of the hydrogen atom as

$$E_n := -\frac{mq^4}{2\hbar^2}\frac{1}{n^2}.$$

This agrees with the empirical formula for the emission and absorption lines of atomic hydrogen which was developed by Johannes Rydberg in 1888.

7.5 Semiclassical Asymptotics

As noted in Section 6.5, the correspondence principle says that one should be able to recover classical physics from quantum mechanics in the *semiclassical* limit where Planck's constant $\hbar$ approaches zero. Although the value of $\hbar$ is fixed in nature, it approaches zero as we change units to longer distance and lower energy scales.

To analyze the semiclassical limit mathematically, we drop the units and simply regard h as a small parameter in the definition of the operator. For Schrödinger operators this means that we write the Hamiltonian as $-h^2\Delta + V$ for $h \to 0$ and study its spectrum in the limit $h \to 0$. We have already seen an example of this in

Carleman's approach to the Weyl asymptotic formula for the Dirichlet Laplacian, which amounted to semiclassical analysis of the operator $-h^2\Delta + 1$.

In this section, we will discuss the problem of determining the semiclassical asymptotics of low-lying eigenvalues of $-h^2\Delta + V$ when V has a single nondegenerate minima. As $h \to 0$, the density of eigenvalues increases, and our physical intuition is that low-lying eigenvalues should be determined by the shape of the potential at the minimum. More precisely, the semiclassical asymptotics of the eigenvalues should correspond to the harmonic oscillator potential defined by the quadratic Taylor approximation to V. This idea was long accepted as folk wisdom in quantum physics, until Barry Simon provided an elegant proof in 1983 [82].

We will assume that V has a single global minimum, which we can place at the origin. The regularity of V away from the origin is not important for the result, but we need enough regularity near the minimum to apply Taylor's theorem. Here is the full set of assumptions on V:

(A1) $V \in L^\infty_{\text{loc}}(\mathbb{R}^n)$ with $V(x) \geq 0$.
(A2) There exists $c > 0$ such that $V(x) \geq c$ outside a compact set.
(A3) V is C^3 in some open neighborhood of $x = 0$.
(A4) V has a unique zero at $x = 0$, which is nondegenerate in the sense that

$$M_{ij} := \frac{\partial^2 V}{\partial x_i \partial x_j}\bigg|_{x=0} \tag{7.50}$$

is a positive definite matrix.

Assumption (A1) guarantees that $-\Delta + V$ is essentially self-adjoint on $C_0^\infty(\mathbb{R}^n)$ by Theorem 7.1. If we split V into components $V_+ + V_-$, where

$$V_+(x) = \max\{V(x), c\}, \qquad V_-(x) = \min\{V(x) - c, 0\},$$

then V_- is compactly supported by (A2). Corollary 7.13 implies that

$$\sigma_{\text{ess}}(-h^2\Delta + V) = \sigma_{\text{ess}}\big(-h^2\Delta + V_+\big).$$

and therefore, since $V_+ \geq c$,

$$\inf \sigma_{\text{ess}}(-h^2\Delta + V) \geq c. \tag{7.51}$$

We are interested in the discrete spectrum,

$$\sigma_{\text{disc}}(-h^2\Delta + V) \subset (0, c).$$

Define the quadratic approximation to V at $x = 0$,

$$V_0(x) := \sum_{i,j=1}^n x_i M_{ij} x_j,$$

where M is the Hessian matrix (7.50), and let

$$P_0 := -\Delta + V_0(x). \tag{7.52}$$

By the computation of the harmonic oscillator spectrum in Theorem 7.4,

$$\sigma(P_0) = \left\{ \sum_{i=1}^{n} (2k_i + 1)\omega_i : k \in \mathbb{N}_0 \right\},$$

where $\omega_1^2, \ldots, \omega_n^2$ denote the eigenvalues of M with $\omega_j > 0$.

Theorem 7.18. *Assume that V satisfies assumptions (A1)–(A4) above. Let $\{\lambda_k(h)\}$ denote the eigenvalues of $-h^2\Delta + V$ and let $\{\mu_k\} = \sigma(P_0)$, where both sequences are arranged in increasing order. For each $m \in \mathbb{N}$, there exists $\varepsilon > 0$ such that for $0 < h < \varepsilon$, $-h^2\Delta + V$ has at least m eigenvalues in the interval $(0, c)$. These eigenvalues satisfy*

$$\lambda_k(h) = h\mu_k + O(h^{\frac{6}{5}})$$

as $h \to 0$.

Proof By assumptions (A3) and (A4) and Taylor's theorem,

$$V(x) = V_0(x) + R(x),$$

where $R(x) = O(x^3)$ for $|x|$ sufficiently small.

To relate $-h^2\Delta + V$ to P_0, we must first scale out the factor h^2. Let U_h be the unitary transformation of $L^2(\mathbb{R}^n)$ given by

$$U_h f(x) := h^{\frac{n}{4}} f(h^{\frac{1}{2}} x).$$

Conjugation by U_h gives

$$U_h(-h^2\Delta + V)U_h^{-1} = -h\Delta + V(h^{\frac{1}{2}} x). \tag{7.53}$$

Let us define $V_h(x) := h^{-1} V(h^{\frac{1}{2}} x)$, and

$$P_h := -\Delta + V_h.$$

This is consistent with the notation used for P_0, in that

$$V_h(x) = V_0(x) + R_h(x),$$

where

$$R_h(x) := h^{-1}R(h^{\frac{1}{2}}x) = O(h^{\frac{1}{2}}x^3), \tag{7.54}$$

for $|h^{\frac{1}{2}}x|$ sufficiently small.

By (7.53), $-h^2\Delta + V$ is unitarily equivalent to hP_h. Thus it suffices to consider the eigenvalues $\{\mu_k(h)\}$ of P_h, for which

$$\lambda_k(h) = h\mu_k(h).$$

Let $\{\phi_k\}$ denote the eigenfunctions of P_0, which take the form

$$\phi_k(x) = q_k(x)e^{-\frac{1}{2}x^t M^{1/2}x}, \tag{7.55}$$

with $q_k(x)$ a polynomial.

To estimate $\mu_k(h)$, we will apply a cutoff to ϕ_k to define an approximate eigenvalue. Let $\chi \in C_0^\infty(\mathbb{R}^n)$ satisfy

$$\chi(x) = \begin{cases} 1, & |x| \le 1, \\ 0, & |x| \ge 2. \end{cases}$$

For $\gamma > 0$, to be chosen later, we set

$$\chi_h(x) := \chi(h^\gamma x).$$

Since $\chi_h(x) = 1$ for $|x| \le h^{-\gamma}$, it follows from (7.55) and the normalization $\|\phi_k\| = 1$ that

$$\left\langle \chi_h\phi_j, \chi_h\phi_k \right\rangle = \delta_{jk} + O(h^\infty), \tag{7.56}$$

with constants that depend on j, k as well as γ.

Our goal is to estimate $\mu_k(h)$ by min–max, using the span of $\{\chi_h\phi_1, \dots, \chi_h\phi_k\}$ as a test subspace. For the numerator of the Rayleigh quotients, we compute

$$\begin{aligned} P_h(\chi_h\phi_k) &= \chi_k(P_0 + R_h)\phi_k - [\Delta, \chi_h]\phi_k, \\ &= \mu_k\chi_h\phi_k + \chi_h R_h\phi_k - [\Delta, \chi_h]\phi_k. \end{aligned} \tag{7.57}$$

The term $[\Delta, \chi_h]\phi_k$ has support in $|x| \ge h^{-\gamma}$, which implies that

$$\left\langle \chi_h\phi_j, [\Delta, \chi_h]\phi_k \right\rangle = O(h^\infty).$$

By (7.54),

$$\|\chi_h R_h\|_\infty = O(h^{\frac{1}{2}-3\gamma}).$$

Therefore, from (7.57) and (7.56) we can estimate

$$\big\langle \chi_h\phi_j, P_h(\chi_h\phi_k)\big\rangle = \mu_k\delta_{jk} + O(h^{\frac{1}{2}-3\gamma}). \tag{7.58}$$

For the subspace $W = \operatorname{span}\{\chi_h\phi_1, \ldots, \chi_h\phi_k\}$, the estimates (7.56) and (7.58) imply that

$$\max_{u\in P_h\setminus\{0\}} \frac{\langle u, P_h u\rangle}{\|u\|^2} = \mu_k + O(h^{\frac{1}{2}-3\gamma}).$$

This holds for all k, with the constant in the order estimate depending on k. By Theorem 5.15, this proves that there exists $\varepsilon > 0$ such that P_h has at least m eigenvalues for $h < \varepsilon$, with

$$\mu_k(h) \le \mu_k + O(h^{\frac{1}{2}-3\gamma}) \tag{7.59}$$

for $k \le m$.

To finish the proof, we need a corresponding lower bound on $\mu_k(h)$. Define $\eta_h \in C^\infty(\mathbb{R}^n)$ so that $\chi_h^2 + \eta_h^2 = 1$. By commuting the cutoffs past the operator, it is easy to check that

$$P_h = \chi_h P_h \chi_h + \eta_h P_h \eta_h - |\nabla\chi_h|^2 - |\nabla\eta_h|^2. \tag{7.60}$$

Multiplication by a smooth bounded function preserves $\mathcal{D}(P_h)$, by the characterization of the domain in Theorem 7.2. Therefore, the differential identity (7.60) implies that

$$\langle u, P_h u\rangle = \langle u, \chi_h P_h \chi_h u\rangle + \langle u, \eta_h P_h \eta_h u\rangle - \big\langle u, \big(|\nabla\chi_h|^2 + |\nabla\eta_h|^2\big)u\big\rangle, \tag{7.61}$$

for $u \in \mathcal{D}(P_h)$.

For the first term on the right in (7.61), we can use (7.54) to estimate

$$\langle u, \chi_h P_h \chi_h u\rangle = \langle u, \chi_h P_0 \chi_h u\rangle + O(h^{\frac{1}{2}-3\delta}). \tag{7.62}$$

To estimate the second term in (7.61), the positivity of $-\Delta$ gives

$$\begin{aligned}\langle u, \eta_h P_h \eta_h u\rangle &\ge \langle u, \eta_h V_h \eta_h u\rangle \\ &\ge \Big(\inf_{\operatorname{supp}\eta_h} V_h\Big)\|\eta_h u\|^2.\end{aligned}$$

Since $|x| \ge h^{-\delta}$ on the support of η_h,

$$\inf_{\operatorname{supp}\eta_h} V_h = h^{-1} \inf_{|x|\ge h^{1/2-\delta}} V(x).$$

The minimum at 0 is nondegenerate, hence $V(x) \geq c|x|^2$ with $c > 0$ for $|x|$ sufficiently small. Therefore, for h sufficiently small,

$$\inf_{\operatorname{supp} \eta_h} V_h \geq c h^{-2\delta}.$$

Therefore, the second term in (7.61) has a lower bound

$$\langle u, \eta_h P_h \eta_h u \rangle \geq c h^{-2\delta} \|\eta_h u\|^2,$$

for h sufficiently small.

Since $\|\nabla \chi_h\|_\infty$ and $\|\nabla \eta_h\|_\infty$ are $O(h^\delta)$ by construction, the final term in (7.61) is estimated by

$$\left\langle u, \left(|\nabla \chi_h|^2 + |\nabla \eta_h|^2\right) u \right\rangle = O(h^{2\delta}).$$

Combining these estimates in (7.61) gives

$$\langle u, P_h u \rangle \geq \langle u, \chi_h P_0 \chi_h u \rangle + c h^{-2\delta} \|\eta_h u\|^2 - O(h^{\frac{1}{2} - 3\delta}) - O(h^{2\delta}). \tag{7.63}$$

To optimize the error estimates, we now set $\delta = \frac{1}{10}$.

Let Π denote the spectral resolution of P_0, defined as in Section 5.4. We can split the first term in (7.63) as

$$\begin{aligned} \langle u, \chi_h P_0 \chi_h u \rangle &= \left\langle u, \chi_h P_0 \Pi_{[0,\mu_k)} \chi_h u \right\rangle + \left\langle u, \chi_h P_0 \Pi_{[\mu_k,\infty)} \chi_h u \right\rangle \\ &\geq \left\langle u, \chi_h P_0 \Pi_{[0,\mu_k)} \chi_h u \right\rangle + \mu_k \|\chi_h u\|^2. \end{aligned} \tag{7.64}$$

Since $\|\chi_h u\|^2 + \|\eta_h u\|^2 = \|u\|^2$, we can combine (7.64) with (7.63) to deduce that

$$\langle u, P_h u \rangle \geq \mu_k \|u\|^2 + \langle u, \chi_h P_0 \chi_h u \rangle - O(h^{\frac{1}{5}}). \tag{7.65}$$

Now suppose that $W \subset \mathcal{D}(P_h)$ has dimension k. The operator $\chi_h P_0 \Pi_{[0,\mu_k)} \chi_h$ has rank at most $k-1$, since μ_k is the kth eigenvalue of P_0. Therefore, there exists a unit vector $u \in W$ such that

$$\chi_h P_0 \Pi_{[0,\mu_k)} \chi_h u = 0.$$

The estimate (7.65) yields

$$\langle u, P_h u \rangle \geq \mu_k \|u\|^2 - O(h^{\frac{1}{5}}).$$

Thus, for any subspace $W \subset \mathcal{D}(P_h)$ of dimension k,

$$\max_{u \in P_h \setminus \{0\}} \frac{\langle u, P_h u \rangle}{\|u\|^2} \geq \mu_k - O(h^{\frac{1}{5}}).$$

Therefore, by the min–max principle,

$$\mu_k(h) \geq \mu_k - O(h^{\frac{1}{5}}). \tag{7.66}$$

The combination of (7.59) and (7.66) gives

$$\mu_k(h) = \mu_k + O(h^{\frac{1}{5}}),$$

and the claimed result follows since $\lambda_k(h) = h\mu_k(h)$. □

Example 7.19. The conclusion of Theorem 7.18 can be illustrated explicitly using Pöschl–Teller potentials, defined by

$$W_l(x) := -\frac{l(l+1)}{\cosh^2 x},$$

for $l \in \mathbb{N}$. The substitution $z = \tanh x$ reduces the eigenvalue equation for $-\Delta + W_l$ to a Legendre equation. A set of independent L^2 solutions is given by $P_l^m(\tanh x)$, where $P_l^m(z)$, using the special Legendrian functions (7.45), for $m = 1, \ldots, l$. The discrete spectrum of $-\Delta + W$ is $\{-1, \ldots, -l^2\}$.

Now consider the semiclassical Schrödinger operator $-h^2\Delta + \tanh^2 x$, which satisfies the hypotheses of Theorem 7.18, with the harmonic oscillator spectrum $2\mathbb{N} - 1$. In the eigenvalue equation,

$$(-h^2\Delta + \tanh^2 x)\phi = \lambda(h)\phi,$$

let us restrict h so that $h^{-2} = l(l+1)$ for $l \in \mathbb{N}$. Dividing by h^2 then reduces the eigenvalue equation to

$$(-\Delta + W_l)\phi = h^{-2}(\lambda(h) - 1)\phi.$$

Thus, by the Pöschl–Teller calculation,

$$h^{-2}(\lambda_k(h) - 1) = -(l - k + 1)^2,$$

for $k = 1, \ldots l$. In other words,

$$\lambda_k(h) = \frac{(2k-1)l - (k-1)^2}{l(l+1)}.$$

Since $l = h^{-1}(1 + O(h))$ under the restriction made above, this gives

$$\lambda_k(h) = (2k-1)h + O(h^2).$$

For comparison to Figure 7.1, some eigenfunctions of $-h^2\Delta + \tanh^2 x$ are plotted in Figure 7.3 ◊

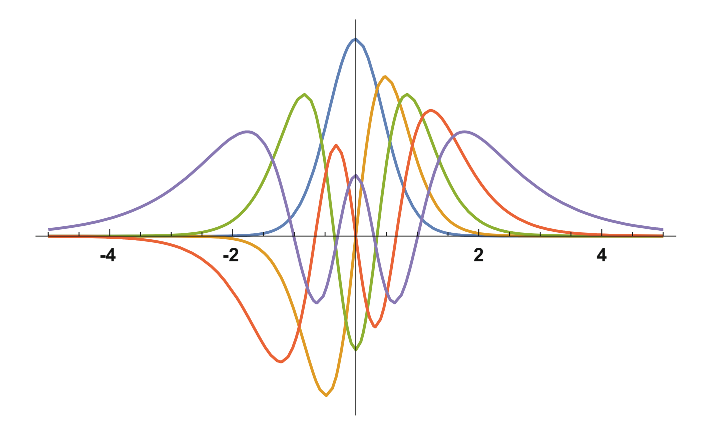


Fig. 7.3 The five eigenfunctions of $-h^2\Delta + \tanh^2 x$ for $h = 1/\sqrt{30}$

7.6 Periodic Potentials

For a continuous periodic function q on $\mathbb{R}$, the ODE

$$-u'' + qu = 0$$

was introduced by George William Hill in 1886 in a study of lunar orbital motion and is called *Hill's equation*. We can apply the analysis of Hill's equation to the spectral theory of the Schrödinger operator $H = -\Delta + V$ on $\mathbb{R}$, where V is a continuous, real-valued, periodic potential. Since V is bounded, H is essentially self-adjoint on $C_0^\infty(\mathbb{R})$ and self-adjoint on $H^2(\mathbb{R})$.

In this section, we will analyze the spectrum of H by means of a theory of periodic systems of ODE developed by Gaston Floquet in 1883.

7.6.1 Floquet Theory

Suppose V is a continuous real-valued function on $\mathbb{R}$ with period π. This choice of period is a standard convention for Hill's equation, inspired by the special case of the Mathieu equation which we will discuss later.

For $\lambda \in \mathbb{R}$, consider the eigenvalue equation for $H := -\Delta + V$,

$$-u'' + Vu = \lambda u. \tag{7.67}$$

By the standard ODE existence result of Picard–Lindelöf (Theorem 9.8), for each $\lambda \in \mathbb{R}$ there exists a C^2 solution to (7.67), which is uniquely determined by the initial conditions $u(0)$ and $u'(0)$. Let us define a pair of solutions u_λ and v_λ by the conditions

$$\begin{cases} u_\lambda(0) = 1, \\ u'_\lambda(0) = 0, \end{cases} \qquad \begin{cases} v_\lambda(0) = 0, \\ v'_\lambda(0) = 1. \end{cases} \tag{7.68}$$

The *Wronskian* of the pair is given by

$$W[u_\lambda, v_\lambda] := u_\lambda v'_\lambda - u'_\lambda v_\lambda.$$

The Wronskian is independent of x, since both functions satisfy (7.67). We can thus compute, by (7.68),

$$W[u_\lambda, v_\lambda] = 1. \tag{7.69}$$

The evolution of this pair of solutions over the course of one period can be expressed in terms of a *monodromy matrix*,

$$M(\lambda) := \begin{pmatrix} u_\lambda(\pi) & v_\lambda(\pi) \\ u'_\lambda(\pi) & v'_\lambda(\pi) \end{pmatrix},$$

which satisfies $\det M(\lambda) = 1$, by (7.69). The trace of the monodromy matrix is called the *discriminant* of H and denoted by

$$\begin{aligned} D(\lambda) &:= \operatorname{tr} M(\lambda) \\ &= u_\lambda(\pi) + v'_\lambda(\pi). \end{aligned}$$

The Picard–Lindelöf theorem implies that the solutions u_λ and v_λ depend smoothly on λ, so $D(\lambda)$ is a smooth function of $\lambda \in \mathbb{R}$. In fact, one can show that $D(\lambda)$ extends to $\lambda \in \mathbb{C}$ as an entire analytic function of order $\frac{1}{2}$. See Magnus and Winkler [61, Thm. 2.2] for a proof.

Example 7.20. In the free case $V = 0$, the solutions u_λ and v_λ are given by

$$u_\lambda(x) = \cos(\sqrt{\lambda}x), \qquad v_\lambda(x) = \frac{\sin(\sqrt{\lambda}x)}{\sqrt{\lambda}},$$

with the convention that $v_\lambda(x) := x$ for $\lambda = 0$. Both u_λ and v_λ are entire functions of $\lambda \in \mathbb{C}$. The monodromy matrix for a period of π is

$$M(\lambda) = \begin{pmatrix} \cos(\pi\sqrt{\lambda}) & -\sin(\pi\sqrt{\lambda})/\sqrt{\lambda} \\ \sqrt{\lambda}\sin(\pi\sqrt{\lambda}) & \cos(\pi\sqrt{\lambda}) \end{pmatrix}.$$

The discriminant is thus the entire function,

$$D(\lambda) = 2\cos(\pi\sqrt{\lambda}).$$

◊

The eigenvalues of the monodromy matrix are called *Floquet multipliers*. The characteristic equation for $M(\lambda)$ is

$$\det(M(\lambda) - t) = t^2 - D(\lambda)t + 1 = 0, \tag{7.70}$$

so in terms of the discriminant the Floquet multipliers are given by

$$\alpha_\pm = \frac{1}{2}\Big[D(\lambda) \pm \sqrt{D(\lambda)^2 - 4}\Big]. \tag{7.71}$$

The following lemma is a special case of Floquet's theorem on periodic systems.

Lemma 7.21. *For $\lambda \in \mathbb{R}$, the solutions of (7.67) are characterized as follows:*

1. *If $|D(\lambda)| < 2$, then the Floquet multipliers are given by $e^{\pm ik}$ for $0 < k < \pi$ and there exist two linearly independent solutions of the form*

$$u_\pm(x) = e^{\pm ikx/\pi} w_\pm(x), \tag{7.72}$$

   where $w_\pm$ are π-periodic functions in $C^2(\mathbb{R})$.
2. *If $|D(\lambda)| > 2$, then the Floquet multipliers are given by $e^{\pm\kappa}$ for $\kappa > 0$ and there exist two linearly independent solutions of the form*

$$u_\pm(x) = e^{\pm \kappa x/\pi} w_\pm(x), \tag{7.73}$$

   where $w_\pm$ are π-periodic functions in $C^2(\mathbb{R})$.
3. *If $|D(\lambda)| = 2$, then the Floquet multipliers are degenerate and given by e^{ik} with $k = 0$ or π. If $M(\lambda)$ is diagonalizable, then there exist two linearly independent solutions of the form (7.72). Otherwise, there exist independent solutions of the form,*

$$u_1(x) = e^{ikx/\pi} w_1(x), \qquad u_2(x) = e^{ikx/\pi}\Big(w_2(x) + \frac{x}{\pi} w_1(x)\Big),$$

   where w_1, w_2 are π-periodic functions in $C^2(\mathbb{R})$.

Proof If $|D(\lambda)| < 2$, then the roots have modulus one and can be written as $\alpha_\pm = e^{\pm ik}$ for $k > 0$. If $(a_\pm, b_\pm)$ denote the eigenvectors in $\mathbb{C}^2$ corresponding to $\alpha_\pm$, let us set

$$u_\pm := a_\pm u_\lambda + b_\pm v_\lambda.$$

By the definition of the monodromy and the fact that $(a_\pm, b_\pm)$ are its eigenvectors,

$$\begin{aligned}\begin{pmatrix}u_\pm(\pi)\\ u'_\pm(\pi)\end{pmatrix} &= M(\lambda)\begin{pmatrix}a_\pm\\ b_\pm\end{pmatrix}\\ &= e^{\pm ik}\begin{pmatrix}a_\pm\\ b_\pm\end{pmatrix}.\end{aligned}$$

The solutions $u_\pm(x+\pi)$ and $e^{\pm ik}u_\pm(x)$ satisfy the same initial conditions at $x = 0$. Therefore, by uniqueness of solutions,

$$u_\pm(x+\pi) = e^{\pm ik}u_\pm(x), \tag{7.74}$$

for all $x \in \mathbb{R}$. This means that the functions,

$$w_\pm(x) := e^{\mp ikx/\pi}u_\pm(x),$$

are periodic, which proves (7.72).

A very similar analysis applies to other cases where $M(\lambda)$ is diagonalizable. If $|D(\lambda)| = 2$, then the eigenvalues are degenerate and equal to ± 1. For $|D(\lambda)| > 2$, the eigenvalues have the form $e^{\pm\kappa}$.

If $|D(\lambda)| = 2$ and $M(\lambda)$ is not diagonalizable, then by putting $M(\lambda)$ into Jordan block form, we can find a pair of solutions u_1, u_2 which satisfy

$$u_1(x+\pi) = e^{ik}u_1(x), \qquad u_2(x+\pi) = e^{ik}\big(u_1(x) + u_2(x)\big).$$

After writing these in terms of w_1, w_2 as above, we can deduce that w_1 and w_2 are periodic. □

Floquet's theorem was rediscovered independently by physicist Felix Bloch in 1928, in the context of electron crystallography. In that application, the quasiperiodic solutions (7.72) are called *Bloch waves*.

7.6.2 Spectrum of H

From the characterization of solutions of the eigenvalue equation in Lemma 7.21, we can now deduce the spectrum of the periodic quantum Hamiltonian in terms of the values of the discriminant.

Theorem 7.22. *For the operator $H = -\Delta + V$ acting on $L^2(\mathbb{R})$, with V continuous and 2π-periodic, let $D(\lambda)$ be the discriminant defined in Section 7.6.1. The spectrum of H is continuous and given by*

$$\sigma(H) = \overline{\{\lambda \in \mathbb{R} : |D(\lambda)| < 2\}}.$$

Proof By Lemma 7.21, solutions of the eigenvalue equation are either quasiperiodic or approach infinity as $|x| \to \infty$ in at least one direction. Hence, H has no L^2 eigenfunctions and the point spectrum is empty.

For $|D(\lambda)| < 2$ we can construct approximate eigenfunctions from the Bloch wave solution u_+ defined by (7.72). Fix $h \in C^\infty(\mathbb{R})$ so that

$$h(t) = \begin{cases} 0, & t \le 0, \\ 1, & t \ge 1. \end{cases}$$

For $j \in \mathbb{N}$, define a sequence of cutoffs $\chi_j \in C_0^\infty(\mathbb{R})$ by

$$\chi_j(x) := h(\pi j - |x|),$$

so that $\chi_j = 1$ on $[-\pi j, \pi j]$. Consider the cutoff sequence $v_j := \chi_j u_+$. Since $u_+(x) = e^{ikx/\pi} w_+(x)$ with w_+ periodic, we can estimate from below,

$$\|v_j\|^2_{L^2(\mathbb{R})} \ge 2j \|w_+\|^2_{L^2(0,\pi)}. \tag{7.75}$$

On the other hand, since u_+ satisfies the eigenvalue equation,

$$\begin{aligned} (H - \lambda)v_j &= -[\Delta, \chi_j]u_+ \\ &= -2\chi_j' u_+' - \chi_j'' u_+. \end{aligned} \tag{7.76}$$

By construction, the derivatives of χ_j are bounded independently of j. Since u_+ is quasiperiodic, we can thus derive from (7.76) the bound

$$\|(H - \lambda)v_j\| \le C \|w_+\|_{H^1(0,\pi)}, \tag{7.77}$$

with C independent of j. Together, (7.75) and (7.77) show that

$$\lim_{j \to \infty} \frac{\|(H - \lambda)v_j\|}{\|v_j\|} = 0.$$

By Theorem 4.16, this implies that $\lambda \in \sigma(H)$.

Finally, for λ such that $|D(\lambda)| > 2$, we will show that $\lambda \in \rho(H)$ by constructing an inverse for $H - \lambda$. The solutions satisfy $u_\pm(x) = e^{\pm\kappa x/\pi} w_\pm(x)$ with $\kappa > 0$ and $w_\pm(x)$ periodic. For $f \in C_0^\infty(\mathbb{R})$, the standard ODE method of variation of parameters yields a solution formula for the equation $(H - \lambda)u = f$:

$$u(x) = \int_{-\infty}^{\infty} K(x, y) f(y)\, dy,$$

where, in terms of the Wronskian of u_- and u_+,

$$K(x,y) := \frac{1}{W[u_-,u_+]}\begin{cases} u_+(x)u_-(y), & x \le y, \\ u_-(x)u_+(y), & x \ge y. \end{cases} \tag{7.78}$$

By construction, $K(x,y)$ is the integral kernel of $(H-\lambda)^{-1}$. Using the exponential decay of $u_\pm$ as $x \to \mp\infty$, we can estimate

$$\int_{-\infty}^{\infty} |K(x,y)|\,dy \le C \int_{-\infty}^{\infty} e^{-\kappa|x-y|/\pi}\,dy = O(1),$$

uniformly in x. The same estimate holds for the integral over x, by symmetry. The Schur test (Exercise 2.11) then shows that $(H-\lambda)^{-1}$ is a bounded operator on $L^2(\mathbb{R})$, proving that $\lambda \in \rho(H)$. □

Example 7.23. In the special case $V(x) = 2b\cos(2x)$, with $b \in \mathbb{R}$, the eigenvalue equation for $-\Delta + V$ is called the *Mathieu equation*,

$$u'' + (\lambda - 2b\cos(x))u = 0. \tag{7.79}$$

Émile Mathieu first derived this equation in the context of the Dirichlet eigenvalue problem for an elliptical domain. The standard solutions of (7.79) are special functions called the "Mathieu sine" and "Mathieu cosine," denoted by $S(\lambda,b,x)$ and $C(\lambda,b,x)$, respectively. In terms of these functions, the discriminant is given by

$$D(\lambda) = \frac{C(\lambda,b,\pi)}{C(\lambda,b,0)} + \frac{S'(\lambda,b,\pi)}{S'(\lambda,b,0)}.$$

Figure 7.4 shows a sample plot of $D(\lambda)$, with spectral bands corresponding to $|D(\lambda)| < 2$. ◊

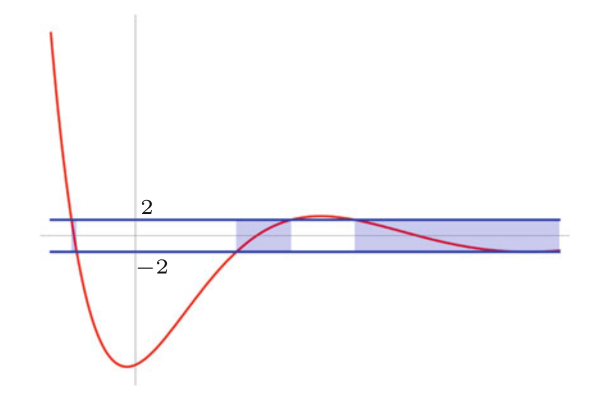


Fig. 7.4 A plot of the discriminant of the Mathieu equation for $b = 2$, with the spectral bands of $-\Delta + 4\cos(2x)$ shaded

The result of Theorem 7.22 can be interpreted in terms of Bloch eigenfunctions defined on the cell $(0, \pi)$. For $\theta \in \mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$, let $-\Delta_\theta$ denote the self-adjoint extension of $-\Delta$ on $(0, \pi)$ defined by the quasiperiodic boundary conditions,

$$f(\pi) = e^{i\theta} f(0), \qquad f'(\pi) = e^{i\theta} f'(0). \tag{7.80}$$

As a self-adjoint operator on a compact interval, $-\Delta_\theta + V$ has a discrete spectrum consisting of real eigenvalues.

To compute these eigenvalues, note that all solutions of the eigenvalue equation have the form $\phi = au_\lambda + bv_\lambda$ for some constants a, b. This function satisfies the quasiperiodic boundary conditions (7.80) provided that (a, b) is an eigenvector of $M(\lambda)$ with eigenvalue $e^{i\theta}$. In other words, λ is an eigenvalue of $-\Delta_\theta + V$ if and only if $e^{i\theta}$ is a Floquet multiplier of the corresponding Hill equation. By (7.71), the Floquet multipliers have modulus one precisely when $|D(\lambda)| \le 2$. Thus, we can restate the conclusion of Theorem 7.22 in the form,

$$\sigma(H) = \bigcup_{\theta \in \mathbb{T}} \sigma(-\Delta_\theta + V). \tag{7.81}$$

It is possible to establish (7.81) more directly via a continuous version of the Floquet transform introduced in Section 4.1.4. For $f \in C_0^\infty(\mathbb{R})$, define

$$Uf(\theta; y) := \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} e^{-ik\theta} f(x + \pi k).$$

This operator extends to a unitary map from $U : L^2(\mathbb{R}) \to L^2(\mathbb{T}; L^2(0, \pi))$. The space $L^2(\mathbb{T}; L^2(0, \pi))$ can be understood in terms of a notion called the "direct integral" (see, e.g., Reed and Simon for details [71, §XIII.16]). The Floquet transform allows us to interpret an operator on $L^2(\mathbb{T}; L^2(0, 2\pi))$ as a family of operators on $L^2(0, 2\pi)$ parametrized by $\theta \in \mathbb{T}$. The conjugate of H by U is the family $-\Delta_\theta + V$, and the characterization (7.81) follows from unitary equivalence.

The Floquet transform, in both discrete and continuous variants, can be adapted to handle more general discrete group actions in higher dimensions. It plays a fundamental role in spectral analysis for electron crystallography.

7.7 Exercises

7.1. For $f \in H^1(\mathbb{R})$ and $\alpha > 0$, consider the quadratic form

$$Q_\alpha[f] := \|f'\|_2^2 - \alpha |f(0)|^2.$$

(a) Show that Q_α is semibounded from below, meaning that there exists a constant c such that

$$Q_\alpha[f] \geq -c\|f\|^2$$

for all $f \in H^1(\mathbb{R})$.

(b) Show that there exists a self-adjoint operator A corresponding to Q_α, such that

$$Q_\alpha[u] = \langle u, Au \rangle$$

for $u \in \mathcal{D}(A)$. Describe $\mathcal{D}(A)$ and the action of A explicitly. (A is interpreted as a Schrödinger operator with potential given by the Dirac point measure $V = -\alpha\delta_0$.)

(c) Show that A has essential spectrum $[0, \infty)$ and discrete spectrum consisting of a single negative eigenvalue.

7.2. Suppose that $H = -\Delta + V$ is a self-adjoint Schrödinger operator on $\mathbb{R}^n$, with domain satisfying $\mathcal{D}(H)$ contained in $H^2_{\mathrm{loc}}(\mathbb{R}^n)$. If ψ is an eigenfunction of H and the potential V is C^m on some open set $U \subset \mathbb{R}^n$, prove that ψ is C^k on U for $k < m + 2 - n/2$.

7.3. Let $V \in L^2_{\mathrm{loc}}(\mathbb{R})$ be a real-valued potential, such that $V(x) \to 0$ as $|x| \to \infty$. This implies $\sigma_{\mathrm{ess}}(-\Delta + V) = [0, \infty)$ by Theorem 7.16. Assume that $V \leq 0$ and is strictly negative on a set of positive measure. Use the min–max principle (Theorem 5.15) to prove that $-\Delta + V$ has at least one strictly negative eigenvalue.

7.4. Let $H = -\Delta + V$ for $V \in C_0^\infty(\mathbb{R}^n)$ with $n \leq 3$, and assume that $V \leq 0$.

(a) Set $V = -\rho^2$ for $\rho \in C_0^\infty(\mathbb{R}^n)$ and define the operator

$$B(\lambda) := \rho(-\Delta - \lambda)^{-1}\rho,$$

which is Hilbert–Schmidt for $\lambda < 0$ by Theorem 7.16. Show that $\lambda < 0$ is an eigenvalue of H if and only if 1 is and eigenvalue of $B(\lambda)$).

(b) Use min–max to show that the eigenvalues of $B(\lambda)$ are continuous increasing functions of λ.

(c) If $n = 1$ or 2, prove that $-\Delta + V$ has at least one negative eigenvalue by studying the behavior of $\|B(\lambda)\|$ as $\lambda \to -\infty$ or 0. [Hint: Use the Fourier transform to estimate $\langle u, B(\lambda)u \rangle$.]

(d) For $n = 3$, show that $B(0)$ is a Hilbert–Schmidt operator and that the number of eigenvalues of H is equal to the number of eigenvalues of $B(0)$ strictly greater than one.

(e) For $n = 3$, use the Hilbert–Schmidt norm to deduce the Birman–Schwinger bound,

$$\#\sigma_{\text{disc}}(H) \le \|B(\lambda)\|_{\text{HS}}^2 := \frac{1}{(4\pi)^2} \int_{\mathbb{R}^3\times\mathbb{R}^3} \frac{V(x)V(y)}{|x-y|^2}\, d^3x\, d^3y.$$

(Note that this shows that H has no eigenvalues if V is sufficiently small, in contrast to (c).)

7.5. If the positions of the two electrons of the helium atom are represented by coordinates $(x, y) \subset \mathbb{R}^3 \times \mathbb{R}^3$, then the quantum Hamiltonian (with units scaled out) takes the form

$$H = -\Delta_x - \Delta_y - \frac{2}{|x|} - \frac{2}{|y|} + \frac{1}{|x-y|},$$

acting on $L^2(\mathbb{R}^6)$.

(a) Use the Kato–Rellich theorem to prove that H is essentially self-adjoint on $C_0^\infty(\mathbb{R}^6)$. [Hint: relative boundedness estimates for the three potentials may be considered separately.]
(b) Based on the hydrogen atom calculation, a reasonable guess for the lowest helium eigenfunction is

$$\psi(x, y) := \frac{a^3}{\pi} e^{-a(|x|+|y|)},$$

with $a > 0$ a free parameter. Estimate λ_1 by minimizing the Rayleigh–Ritz quotient of ψ over a. (For comparison, the precise value is $\lambda_1 \doteq -1.452$ in these units.)

7.6. Suppose ψ is a real-valued eigenfunction of a Schrödinger operator $-\Delta + V$ on $\mathbb{R}$, with eigenvalue λ. The set $\{V > \lambda\}$ is called the "classically forbidden" region. (A classical particle cannot exist in this set, because kinetic energy cannot be negative.) Suppose $V > \lambda$ on (x_0, ∞), and that ψ is at least C^2 on this interval. Prove the exponential decay formula,

$$|\psi(x)| \le \left(|\psi(x_0)| + c_0\sqrt{x - x_0}\right) \exp\left(-\int_{x_0}^{x} \sqrt{V(t) - \lambda}\, dt\right)$$

for $x > x_0$, where $c_0 := \sqrt{|\psi(x_0)\psi'(x_0)|}$, using the following steps.

(a) By changing sign if needed, we can assume that $\psi(x_0) \ge 0$. Show that this implies that $\psi(x) > 0$ and $\psi'(x) < 0$ for all $x \ge x_0$.

(b) For $x \geq x_0$, let

$$h(x) := \exp\left(\int_{x_0}^{x} \sqrt{V(t) - \lambda}\, dt\right).$$

Show that $(h^2\psi\psi')' = [(h\psi)']^2$, and use this to prove that

$$\int_{x_0}^{x} \left[(h\psi)'\right]^2 dt \leq c_0^2.$$

(c) Use (b) to estimate $\int_{x_0}^{x} (h\psi)'\, dt$.

7.7. Consider the quantum harmonic oscillator $H = -\Delta + |x|^2$ in $\mathbb{R}^3$, whose spectrum is given by Theorem 7.4. Use spherical coordinates as in Section 7.4 to derive the eigenvalues and express the eigenfunctions in terms of spherical harmonics.

Notes

Because of their importance in quantum mechanics, Schrödinger operators are one of the most thoroughly studied areas of spectral theory. The literature is vast, and we have given only a few introductory results in this chapter. For additional background on the spectral theory of Schrödinger operators, see Davies [24], Edmunds and Evans [28], Hislop and Segal [44], and Reed and Simon [71]. For a mathematically oriented introduction to quantum mechanics, see Gustafson and Sigal [40] or Hall [41].

The issue of dynamics, i.e., the evolution of solutions of the full Schrödinger equation (7.11) as a function of time, is closely related to the theory of one-parameter groups mentioned in the notes to Chapter 5. This theory leads naturally to the subject of scattering theory, a central topic in quantum mechanics which we have not touched on here. For mathematical treatments of scattering theory, see for example Hislop and Segal [44], Reed and Simon [72], and Yafaev [96].

The perturbative methods discussed in Sections 7.2 and 7.3 are developed in greater detail in Kato [49] and Reed and Simon [70, Ch. X].

There is a huge literature on the semiclassical theory of Schrödinger operators. Most results in semiclassical analysis rely on machinery (pseudodifferential operator calculus) that lies outside the scope of this book. For an introduction to semiclassical techniques, see Zworski [97].

Chapter 8
Operators on Graphs



The spectrum of a finite graph can be defined in terms of its adjacency matrix. In "spectral graph theory," the eigenvalues of this matrix (or related matrices) are used to analyze properties of the graph, such as connectivity.

The integer lattice discussed in Section 4.1.4 could be thought of as an infinite discrete graph, with lattice points as vertices and edges connecting nearest neighbors. The discrete Laplacian on $\mathbb{Z}^n$, as defined in (4.10), involves only neighboring vertices, and in fact differs from the adjacency matrix by a multiple of the identity. We can thus adapt the formula for $\Delta_{\mathbb{Z}^n}$ to define the discrete Laplacian on an arbitrary graph.

Spectral graph theory has a long history as an important tool in combinatorics, with applications to network theory and computer science, as well as number theory, chemistry, and mathematical physics.

Another way to build spectral models from graphs is by considering metric graphs, for which a length is assigned to each edge. This identifies the edges with intervals in $\mathbb{R}$, allowing us to consider differential operators acting on a Hilbert space defined as the direct sum of the L^2 spaces for each edge interval. In physical applications, the operator is usually taken to be either the one-dimensional Laplacian or a Schrödinger operator. This combination of a metric graph equipped with a quantum Hamiltonian is called a *quantum graph*. Research in quantum graphs has been motivated by in large part by physical applications, such as understanding the electromagnetic properties of carbon nanostructures.

Note that the key difference between the discrete and continuous cases lies in the support of the functions. In the discrete case the Hilbert space consists of functions on the vertices, while in the quantum graph case they live on the edges. It is possible to synthesize these two approaches and consider functions taking values on both edges and vertices.

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1_8

In this chapter we will give a very brief introduction to the spectral theory of graphs, focusing mainly on the case of quantum graphs. Our goal here is to highlight the connections to the spectral theory developed elsewhere in this book. We refer the reader to the notes at the end of the chapter for more detailed background sources.

8.1 Combinatorial Laplacians

The discrete Laplacian operators discussed in Section 4.1.4 involved a sum over nearest neighbors in the lattice $\mathbb{Z}^n$. This definition extends naturally to the framework of graphs, with vertices considered to be neighbors if they are joined by an edge.

Let Γ be a graph with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$. Multiple edges between the same pair of vertices are allowed. In principle we could allow loops also, but these would be irrelevant for reasons explained below. For two vertices $v_1, v_2 \in \mathcal{V}$, we write $v_1 \sim v_2$ if v_1 and v_2 are connected by an edge of Γ. We will assume that the maximum degree of Γ is finite, i.e., the number of edges connected to a vertex is bounded.

Taking the definition of $-\Delta_{\mathbb{Z}^n}$ in (4.10) as a model, we define the *combinatorial Laplacian* of Γ as the operator on $\ell^2(\mathcal{V})$ given by

$$Lf(v_i) = \sum_{v_j \in \mathcal{V}} A_{ij}(f(v_i) - f(v_j)), \tag{8.1}$$

where A_{ij} is the *adjacency matrix*, which counts the number of (undirected) edges connecting v_i to v_j. This operator is associated with the quadratic form

$$Q[f, f] := \sum_{v_i, v_j \in \mathcal{V}} A_{ij} |f(v_i) - f(v_j)|^2. \tag{8.2}$$

The sums in (8.1) and (8.2) are finite, by the assumption that Γ has finite maximum degree. Note that if we did allow loops in Γ, their contribution to these sums would be zero, so henceforth we assume that Γ has no loops. More general operators could be defined by attaching weights to the edges and/or vertices, but we will consider those cases here.

The adjacency matrix is symmetric, so if the maximal degree of a vertex of Γ is finite, then $Q[\cdot, \cdot]$ is a bounded symmetric form, and L is a bounded self-adjoint operator acting on $\ell^2(\mathcal{V})$.

If Γ is a finite graph with m vertices, then $\ell^2(\mathcal{V})$ is naturally isomorphic to $\mathbb{C}^m$, and L is represented by a symmetric $m \times m$ matrix. We can write this matrix as

$$L = D - A,$$

where A is the adjacency matrix defined above, and D is the diagonal matrix with entries given by the degree of each vertex. For example, the graph pictured in Figure 8.1 has

$$D = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

yielding the combinatorial Laplacian

$$L = \begin{pmatrix} 3 & -2 & -1 & 0 \\ -2 & 4 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}.$$

For infinite graphs, L might have continuous spectrum, as demonstrated by the examples in Section 4.1.4.

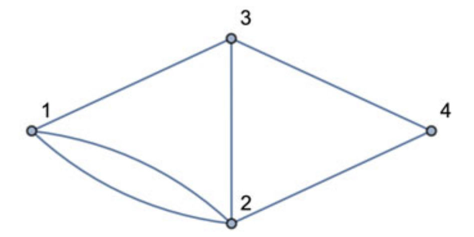


Fig. 8.1 Graph with four vertices and six edges

When the graph is not finite, we at least have that L is a symmetric operator the space $C_0(\mathcal{V})$ consisting of functions supported on finitely many vertices. This operator admits a unique self-adjoint extension by the following:

Lemma 8.1. *As an operator on $\ell^2(\mathcal{V})$, L is essentially self-adjoint on the domain $C_0(\mathcal{V})$.*

Proof Since Lf is well defined for any function on $\mathcal{V}$, we can characterize the adjoint domain by

$$\mathcal{D}(L^*) = \left\{ f \in \ell^2(\mathcal{V}) : \ Lf \in \ell^2(\mathcal{V}) \right\}.$$

Because L is positive, to prove essential self-adjointness it suffices to show that $L^* + 1$ is injective, by Exercise 3.11.

Suppose $f \in \mathcal{D}(L^*)$ satisfies $(L^* + 1)f = 0$. This means precisely that $f \in \ell^2(\mathcal{V})$ and

$$Lf = -f. \tag{8.3}$$

If $f(w_0) > 0$ at some vertex $w_0 \in V$, then evaluating (8.3) at w_0 gives

$$\sum_{v \in \mathcal{V}} A_{v,w_0}(f(w_0) - f(v)) < 0.$$

This means that w_0 has at least one neighbor, call it w_1, for which $f(w_1) > f(w_0)$. Repeating this process inductively gives a sequence of neighboring vertices $\{w_j\}$ such that the sequence of values $f(w_j)$ is strictly increasing. Since Γ is infinite, this would contradict the assumption that $f \in \ell^2(\mathcal{V})$. Therefore $f \leq 0$ at all vertices.

The same argument applies if we start from $f(w_0) < 0$, so $(L^* + 1)f = 0$ implies that $f = 0$. Therefore, L is essentially self-adjoint. □

Much of the interest in combinatorial spectral theory lies in the connections between $\sigma(L)$ and the geometric features of Γ. We will illustrate this idea with a few basic results for finite graphs.

For a finite graph Γ with m vertices, let us arrange the eigenvalues in increasing order,

$$0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m,$$

with λ_1 corresponding to the constant eigenfunction. Some basic properties of the eigenvalues can be deduced by taking the trace of L. If Γ has q edges (assuming no loops), then the sum of degrees of the vertices equals $2q$. Therefore, since $\operatorname{tr} A = 0$,

$$\sum_{j=2}^{m} \lambda_j = \operatorname{tr} D = 2q.$$

It follows that

$$\lambda_2 \leq \frac{2q}{m-1}.$$

This seemingly crude estimate is sharp in the case of the complete graph with m vertices, for which $\lambda_2 = \cdots = \lambda_m = m$.

Eigenvalues can also be estimated through the max–min principle (5.31), which says, for example, that

$$\lambda_2 = \inf_{f \perp 1} \frac{Q[f, f]}{\|f\|^2}. \tag{8.4}$$

As an application, let us derive a basic lower bound for λ_2 in terms of the diameter of the graph. The (discrete) distance between two vertices is the minimal number of edges in a path connecting them, and the *diameter* of Γ is the maximum distance between two vertices.

Theorem 8.2. *For a finite graph of diameter d with m vertices,*

$$\lambda_2 \geq \frac{1}{md}.$$

Proof Consider a real-valued eigenfunction ϕ_2 for λ_2. By (8.4), ϕ_2 satisfies the orthogonality condition

$$\sum_{v \in \mathcal{V}} \phi_2(v) = 0, \tag{8.5}$$

as well as

$$\lambda_2 = \frac{Q[\phi_2, \phi_2]}{\|\phi_2\|^2}. \tag{8.6}$$

Let v_0 denote a vertex where $|\phi_2|$ attains its maximum value. A crude estimate gives

$$\|\phi_2\|^2 \leq m\phi_2(v_0)^2. \tag{8.7}$$

By (8.5), there is at least one vertex v_k where has sign opposite to $\phi_2(v_0)$. We can thus find a sequence of adjacent vertices $v_1, \ldots, v_k$, with $k \leq d$, such that

$$|\phi_2(v_0)| \leq |\phi_2(v_0) - \phi_2(v_k)|. \tag{8.8}$$

By writing $\phi_2(v_0) - \phi_2(v_k)$ as a telescoping sum and applying Cauchy–Schwarz, we can estimate

$$\begin{aligned}
(\phi_2(v_0) - \phi_2(v_k))^2 &= \left(\sum_{j=1}^{k} \phi_2(v_j) - \phi_2(v_{j-1}) \right)^2 \\
&\leq k \sum_{j=1}^{k} \big(\phi_2(v_j) - \phi_2(v_{j-1})\big)^2 \\
&\leq k Q[\phi_2, \phi_2].
\end{aligned}$$

Thus from (8.8) we have

$$Q[\phi_2, \phi_2] \geq \frac{1}{d}\phi_2(v_0)^2. \tag{8.9}$$

Applying (8.9) and (8.7) to (8.6) yields the lower bound on λ_2. □

8.2 Quantum Graphs

A *metric graph* is a locally finite graph Γ with vertices $\mathcal{V}$ and edges $\mathcal{E}$, equipped with a length function $\ell : \mathcal{E} \to \mathbb{R}_+$. Each edge $e \in \mathcal{E}$ of a metric graph can be identified with an interval $[0, \ell(e)]$, at least up to orientation. The edges are undirected, meaning that neither orientation is preferred. Loops and multiple edges between a pair of vertices are allowed. Although open edges identified with $[0, \infty)$ are sometimes allowed in the definition, we assume here that ℓ is finite-valued.

A function f on Γ is defined as a collection of functions $f_e : [0, \ell(e)] \to \mathbb{C}$ for each edge $e \in \mathcal{E}$. The Hilbert space associated with the metric graph Γ is thus defined as

$$L^2(\Gamma) := \bigoplus_{e\in\mathcal{E}} L^2(0, \ell(e)).$$

A *quantum graph* is a metric graph equipped with a quantum mechanical Hamiltonian operator H. For simplicity, we will focus on the case where this operator is the Laplacian, $-\Delta$, which acts as the differential operator $-\partial_x^2$ on each edge.

The extension of $-\Delta$ to a self-adjoint operator on $L^2(\Gamma)$ requires the specification of a domain. One possible choice is to apply Dirichlet boundary conditions on each edge, so that the functions vanish at each vertex. Although this clearly defines a self-adjoint operator, it has the effect of decoupling the edges. With Dirichlet conditions at all vertices, the graph structure is forgotten and Γ is effectively a disjoint union of intervals.

The most common assumption in quantum graph theory is the Neumann–Kirchoff vertex condition, which is a natural analog of Neumann conditions on an interval. We saw in Section 6.1.3 that the Neumann Laplacian for a bounded open set can be derived from the H^1 inner product, without imposing restrictions explicitly. We can take the same approach for quantum graphs. From Example 2.24 we can see that functions in $H^1(0, \ell)$ are continuous on the closed interval $[0, \ell]$. We extend this definition to metric graphs by setting

$$H^1(\Gamma) := \Big\{ f \in \bigoplus_{e\in\mathcal{E}} H^1(0, \ell(e)) : \ f \text{ is continuous across vertices} \Big\},$$

with the inner product

$$\|f\|_{H^1}^2 := \sum_{e\in\mathcal{E}} \|f_e\|_{H^1}^2.$$

Following the Neumann definition (6.18), we introduce the domain

$$\mathcal{D}(-\Delta) := \big\{ u \in H^1(\Gamma) : \ \langle u, \cdot \rangle_{H^1} \text{ extends to } L^2(\Gamma) \text{ as a bounded functional} \big\}. \tag{8.10}$$

For $u \in \mathcal{D}(-\Delta)$, the condition on $\langle u, \cdot \rangle_{H^1}$ implies, by the Riesz lemma (Theorem 2.28), that there exists a unique $f \in L^2(\Gamma)$ such that

$$\langle u, g \rangle_{H^1} = \langle f, g \rangle$$

for all $g \in H^1(\Gamma)$. We can then define $-\Delta u$ as $f - u \in L^2(\Gamma)$, so that

$$\langle u', g' \rangle = \langle -\Delta u, g \rangle. \tag{8.11}$$

By the Friedrichs extension construction developed in Section 6.1.2, we obtain the following:

Theorem 8.3. *For a metric graph Γ, the Laplacian is self-adjoint on the domain* (8.10).

The terminology "Neumann–Kirchoff" comes from Gustav Kirchoff's law of electrical circuits, which says that the total current entering a junction point is equal to the total current flowing out of that point. The explanation for the usage of this term for the operator defined by (8.10) is given by the following:

Lemma 8.4. *If $f \in \mathcal{D}(-\Delta)$, then f restricts to function in $C^1[0, \ell(e)]$ on each edge $e \in \mathcal{E}$. Furthermore, the derivatives at each vertex $v \in \mathcal{V}$ satisfy*

$$\sum_{e \ni v} f_e'(v) = 0, \tag{8.12}$$

where $e \ni v$ means e contains v as an endpoint, and the convention is that the derivatives are taken outward from the vertex.

Proof As noted above, $f \in \mathcal{D}(-\Delta)$ implies that Δf exists in the weak sense and is contained in $L^2(\Gamma)$. Thus, each edge component f_e has a weak derivative $f_e' \in H^1(0, \ell(e))$. Functions in $H^1(0, \ell(e))$ extend to H^1 functions on $\mathbb{R}$, by Lemma 6.14, and H^1 functions are continuous in dimension one, by Theorem 2.26. Therefore $f_e' \in C^0[0, \ell(e)]$. By Lemma 2.22, the continuity of f_e' implies that f_e is classically differentiable, and thus $f_e \in C^1[0, \ell(e)]$.

For $f, g \in \mathcal{D}(-\Delta)$ integration by parts on each segment gives

$$\langle f', g' \rangle = \langle -\Delta f, g \rangle + \sum_{v \in \mathcal{V}} \sum_{e \ni v} f_e'(v) g_e(v), \tag{8.13}$$

with derivatives taken outward from the vertex. Since (8.11) holds for all $g \in \mathcal{D}(-\Delta)$, from (8.13) we can deduce that (8.12) holds for each $v \in \mathcal{V}$. □

At a vertex of degree 2, (8.12) reduces to the condition that f' extends continuously across the vertex. Thus, under Neumann–Kirchoff conditions, a pair of edges meeting at a vertex of degree 2 is indistinguishable from a single combined edge with the lengths added together.

Although Neumann–Kirchoff conditions are the most common choice, self-adjointness holds for a more general collection of mixed vertex conditions. See Berkolaiko and Kuchment [11, Thm. 1.4.4] for a description of the full set of possibilities.

8.3 Spectral Properties of Compact Quantum Graphs

A quantum graph is *compact* if it has a finite number of edges, each of finite length. Based on our experience with the Laplacian bounded regions in $\mathbb{R}^n$, we would expect a compact quantum graph to have discrete spectrum. Indeed, this follows from an easy extension of the argument developed in Section 6.2.

Theorem 8.5. *The Laplacian on a compact quantum graph has compact resolvent. Therefore, its spectrum consists of a discrete set of eigenvalues $\{\lambda_j\}$ with $\lambda_j \to \infty$.*

Proof By Rellich's theorem (Theorem 6.9), the embedding $H^1(a,b) \to L^1(a,b)$ is compact for a finite interval $(a,b) \subset \mathbb{R}$. This implies that the embedding

$$\bigoplus_{e\in\mathcal{E}} H^1(0,\ell(e)) \to L^2(\Gamma)$$

is compact, because $\mathcal{E}$ is a finite set by assumption. Since $H^1(\Gamma)$ is a subspace of $\oplus_{e\in\mathcal{E}} H^1(0,\ell(e))$, it follows that the embedding

$$H^1(\Gamma) \to L^2(\Gamma)$$

is compact.

By the same argument used in the proof of Theorem 6.8, the resolvent of $-\Delta$ is bounded as a map $L^2(\Gamma) \to H^1(\Gamma)$. Therefore, the resolvent is compact as an operator on $L^2(\Gamma)$. The structure of the spectrum follows from the Hilbert–Schmidt theorem (Theorem 4.21). □

Under Neumann–Kirchoff conditions, the lowest eigenvalue for any compact graph is $\lambda_1 = 0$, corresponding to the constant functions. The multiplicity of zero as an eigenvalue is the number of connected components of the graph.

If ψ is an eigenfunction with eigenvalue $\lambda = \kappa^2$ for $\kappa > 0$, then on each edge the eigenvalue equation implies that ψ_e is a combination of $e^{\pm i\kappa x}$. Eigenvalues can be computed directly by matching these solutions at the vertices and solving a linear system for the coefficients.

Example 8.6. Consider an equilateral star graph consisting of $m \geq 2$ edges of length ℓ/m, meeting at a central vertex, as illustrated in Figure 8.2. The Neumann conditions on the outer vertex require the eigenfunctions to be proportional to $\cos(\kappa x)$ in each edge, assuming $x = 0$ at the outer vertex. If the coefficient of the jth edge is a_j, then at the central vertex we have the continuity condition,

$$a_j \cos(\kappa\ell/m) \text{ is independent of } j,$$

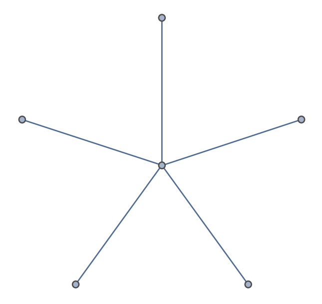


Fig. 8.2 Star graph with five edges

and the Kirchoff circuit condition

$$\sum_{j=1}^{k} a_j \sin(\kappa\ell/m) = 0.$$

There are two families of solutions. Either

$$\cos(\kappa\ell/m) = 0 \quad \text{and} \quad a_1 + \cdots + a_k = 0,$$

or

$$\sin(\kappa\ell/m) = 0 \quad \text{and} \quad a_1 = a_2 = \cdots = a_k.$$

Therefore,

$$\sigma(-\Delta) = \left\{ \left(\frac{\pi m j}{2\ell} \right)^2 : j \in \mathbb{N}_0 \right\},$$

where the multiplicity is $m - 1$ if j is odd and 1 if j is even. $\diamondsuit$

Example 8.7. Let Γ be an equilateral dipole graph, as illustrated in Figure 8.3, consisting of two vertices joined by $m \geq 2$ edges of length ℓ/m. The continuity and vertex conditions for $\lambda = \kappa^2$ admit solutions only when $\sin(\kappa\ell/m) = 0$. For each $\kappa\ell/m \in \pi\mathbb{N}$, then there are $k - 1$ independent eigenfunctions with each edge solution proportional to $\sin(\kappa x)$, and 1 with edge solution $\cos(\kappa x)$. Thus,

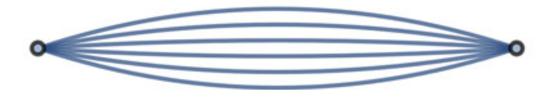


Fig. 8.3 Dipole (or pumpkin) graph with six edges

$$\sigma(-\Delta) = \left\{ \left(\frac{\pi m j}{\ell} \right)^2 : j \in \mathbb{N}_0 \right\}, \tag{8.14}$$

with multiplicity k for all nonzero eigenvalues.

Suppose instead we consider an equilateral flower graph, as illustrated in Figure 8.4. If the flower consists of $m \geq 2$ loops attached to a single vertex, then the only difference from the dipole case is the fact that the cosine edge solutions only occur for $\kappa\ell/m \in 2\pi\mathbb{N}$, by the continuity condition. The eigenvalue set is still given by (8.14), but the multiplicity of $(\pi m j/\ell)^2$ is only $m-1$ for j odd. ♢

The dipole graph exhibits a novel feature of quantum graphs: there exist eigenfunctions that vanish completely on some edges. This means that unique continuation, which we proved for the Laplacian on bounded domains in Corollary 6.16, does not hold for quantum graphs.

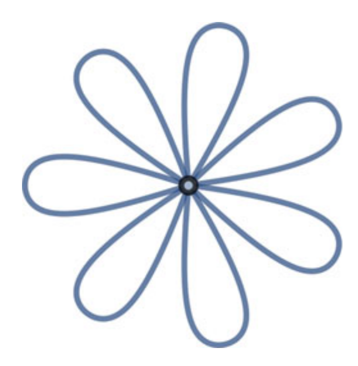


Fig. 8.4 Flower graph with seven edges

8.4 Eigenvalue Comparison

Eigenvalues of a quantum graph may be characterized by the min–max principle of Theorem 5.15. Since the self-adjoint extension of $-\Delta$ is derived from a quadratic form on $H^1(\Gamma)$, we can apply the argument used in Theorem 6.18 to replace the operator domain $\mathcal{D}(-\Delta)$ by the form domain $H^1(\Gamma)$ in the min–max statement. This yields the following:

Theorem 8.8. *Let Γ be a compact quantum graph, with eigenvalues $\{\lambda_k\}$ arranged in increasing order and repeated according to multiplicity. If Λ_k denotes the collection of subspaces of $H^1(\Gamma)$ of dimension k, then*

$$\lambda_k = \min_{W \in \Lambda_k} \left\{ \max_{u \in W \setminus \{0\}} \frac{\|u'\|^2}{\|u\|^2} \right\}. \tag{8.15}$$

As in other cases we have considered, the min–max principle is extremely useful for comparison estimates.

Corollary 8.9. *Suppose that Γ and $\tilde{\Gamma}$ are compact graphs, with respective eigenvalue sets $\{\lambda_k\}$ and $\{\tilde{\lambda}_k\}$, listed in increasing order and repeated according to multiplicity. Assume that the two graphs are related in one of the following ways:*

(a) *Γ is obtained from $\tilde{\Gamma}$ by identifying a pair of vertices, preserving the edge attachments.*
(b) *$\tilde{\Gamma}$ is obtained from Γ by splitting an edge at an interior point, introducing two new vertices.*
(c) *$\tilde{\Gamma}$ is formed from Γ by attaching a new graph to a single vertex of Γ.*
(d) *$\tilde{\Gamma}$ is formed from Γ by lengthening an edge.*

Then, for each $k \in \mathbb{N}$,

$$\lambda_k \geq \tilde{\lambda}_k.$$

Proof By condition (a) there is a natural embedding of $H^1(\Gamma)$ into $H^1(\tilde{\Gamma})$, since both graphs have the same edge sets and the only difference between the H^1 spaces is that fewer continuity constraints are imposed for $H^1(\tilde{\Gamma})$. Since the Rayleigh quotient is preserved under this embedding, the eigenvalue comparison follows immediately from the fact that $\tilde{\Gamma}_k$ is given by minimizing over a larger collection of subspaces.

Introducing a new vertex of degree 2 within an edge does not affect the spectrum. Thus we can regard (b) as a special case of (a).

Now suppose that (c) holds. Let v_0 denote the vertex of Γ where the new graph Γ' is attached to form $\tilde{\Gamma}$. Given a function $u \in H^1(\Gamma)$, we define an extension $\tilde{u} \in H^1(\tilde{\Gamma})$ by

$$\tilde{u} = \begin{cases} u, & \text{on } \Gamma, \\ u(v_0), & \text{on } \Gamma'. \end{cases}$$

Because $\tilde{u}' = 0$ on Γ', the Rayleigh quotient of $\tilde{u}$ is bounded above by that of u. Thus $\tilde{\lambda}_k$ is computed by minimizing a smaller quantity over a larger set, relative to λ_k.

Condition (d) can be handled with a combination of the previous results. First add a new vertex of degree 2 within an edge of Γ, leaving the spectrum unchanged. Then attach a loop to this new vertex, which can only decrease the eigenvalues by (c). Finally, split the original edge at this new vertex to form a single longer edge, which can only decrease the eigenvalues by (a). □

Our first application of Corollary 8.9 gives a lower bound on the eigenvalues depending on the total length of the graph,

$$\ell(\Gamma) := \sum_{e \in \mathcal{E}} \ell(e).$$

The following result is due to Friedlander [32].

Theorem 8.10. *If Γ is a connected compact quantum graph with total length $\ell(\Gamma)$, then*

$$\lambda_k \geq \frac{\pi^2 k^2}{4\ell(\Gamma)^2}. \tag{8.16}$$

Proof Friedlander's argument involves reducing the graph to a tree, using Corollary 8.9(b). To avoid some technicalities, we will prove the result only for k even, using a simpler argument inspired by Kurasov–Naboko [54].

Let $\tilde{\Gamma}$ be the graph obtained from Γ by keeping the same set of vertices and doubling each edge. The doubled edges are assigned the original lengths. A function $u \in H^1(\Gamma)$ can be lifted to a function $\tilde{u} \in H^1(\tilde{\Gamma})$ by simply copying the values of u along repeated edges. Note $\tilde{u}$ has the same Rayleigh quotient as u, because both numerator and denominator are multiplied by 2. It thus follows from Theorem 8.8 that

$$\lambda_k \geq \tilde{\lambda}_k.$$

Note $\tilde{\Gamma}$ is a connected graph for which all vertices have even degree. By the famous result of Euler, $\tilde{\Gamma}$ thus admits a cycle C which visits each edge once. We can think of C as a graph with vertices of degree 2, such that $\tilde{\Gamma}$ can be formed by identifying these vertices in pairs. Therefore, by Corollary 8.9(a),

$$\tilde{\lambda}_k \geq \lambda_k(C).$$

The nonzero eigenvalues of C each have multiplicity 2, so that

$$\lambda_k(C) = \left(\frac{2\pi[k/2]}{\ell(C)}\right)^2.$$

Since $\ell(C) = 2\ell(\Gamma)$, it follows that

$$\lambda_k \geq \left(\frac{\pi[k/2]}{\ell(\Gamma)}\right)^2.$$

This establishes the sharp result for k even, and a suboptimal lower bound of $\pi^2(k-1)^2/4\ell(\Gamma)^2$ when k is odd. □

Friedlander's proof of Theorem 8.10 also shows that equality in (8.16) holds for λ_2 only if Γ is an interval, and for λ_k with $k \geq 3$ only if Γ is a star graph with k edges (Example 8.6).

The example of the star graph shows λ_2 could be arbitrarily large for a fixed value of $\ell(\Gamma)$. However, upper bounds are possible if we include other information such as the number of edges or vertices. To illustrate this, we will derive a simple estimate for λ_2 based on comparison to the following case:

Example 8.11. Let Γ be a flower graph with only two petals, of lengths ℓ_1 and ℓ_2. We claim that

$$\lambda_2 = \frac{4\pi^2}{(\ell_1 + \ell_2)^2}. \tag{8.17}$$

To see this, note first that we can define an eigenfunction with the edge functions,

$$u_1(x) = \cos(\kappa(x - \ell_1/2)), \qquad u_2(x) = \cos(\kappa(x + \ell_1/2)),$$

where $\kappa = 2\pi/(\ell_1 + \ell_2)$. This shows that $\lambda_2 \leq \kappa^2$.

On the other hand, Γ can be constructed from a loop C of length $\ell_1 + \ell_2$ by introducing two appropriately spaced vertices of degree 2 and then identifying them. Corollary 8.9(a) thus implies that $\lambda_2 \geq \lambda_2(C) = \kappa^2$. ♢

Theorem 8.12. *Suppose that Γ is a compact quantum graph with at least two edges. If the two longest edges of Γ have lengths ℓ_1 and ℓ_2, then*

$$\lambda_2 \leq \frac{4\pi^2}{(\ell_1 + \ell_2)^2}.$$

In particular, $\lambda_2 \leq (\pi/\bar{\ell})^2$, where $\bar{\ell}$ denotes the mean edge length of Γ

Proof If we identify all of the vertices of Γ to form a flower graph, this can only raise the eigenvalues by Corollary 8.9(a). Dropping petals from this flower can also only raise the eigenvalues, by Corollary 8.9(c). Therefore, we have

$$\lambda_k \leq \lambda_k(\Gamma_{\ell_1,\ell_2}),$$

where Γ_{ℓ_1,ℓ_2} is the flower with two petals of lengths ℓ_1, ℓ_2. The estimate now follows from (8.17). □

8.5 Eigenvalue Asymptotics

By studying the linear system for coefficients of eigenfunction components in a compact quantum graph, we can develop the analog of a characteristic polynomial, whose roots give the eigenvalues. The presentation in this section is based on the more general discussion in Berkolaiko and Kuchment [11, §2.1]. Here we consider only Neumann–Kirchoff vertex conditions for simplicity.

As we noted above, the lowest eigenvalue of a compact graph is zero, with multiplicity equal to the number of connected components of the graph. Thus we are only concerned with eigenvalues $\lambda = \kappa^2 > 0$, for which the eigenfunctions are linear combinations of $e^{\pm i\kappa x}$ in each edge.

Let us focus first on a single vertex v, with d edges. Assume that these are parametrized by variables $x_1, \ldots, x_d$ such that $x_j = 0$ at v. In terms of these coordinates, an eigenfunction ψ with eigenvalue κ^2 takes the local form,

$$\psi_j(x_j) = \alpha_j e^{i\kappa x_j} + \beta_j e^{-i\kappa x_j} \tag{8.18}$$

on the jth edge, for some complex coefficients α_j and β_j. We label the terms $e^{i\kappa x_j}$ as "outgoing" from the vertex, while the $e^{-i\kappa x_j}$ solutions are "incoming." This distinction reflects the behavior of wavefronts of the corresponding solutions of the wave equation.

The continuity requirement at the vertex implies that

$$\psi(v) = \alpha_j + \beta_j, \tag{8.19}$$

independent of j. After factoring out $i\kappa$, the Kirchoff circuit condition is

$$\sum_{j=1}^{m} (\alpha_j - \beta_j) = 0. \tag{8.20}$$

If Q denotes the orthogonal projection onto the span of $(1, \ldots, 1)$ in $\mathbb{C}^d$, then in terms of vectors $\alpha = (\alpha_1, \ldots, \alpha_d)$ and $\beta = (\beta_1, \ldots, \beta_d)$, the conditions (8.19) and (8.20) take the form

$$Q(\alpha - \beta) = 0, \qquad (I - Q)(\alpha + \beta) = 0.$$

Adding these equations together gives

$$\alpha = (2Q - I)\beta. \tag{8.21}$$

This is called the *scattering relation* for the vertex; it describes the relation between incoming and outgoing solutions. For future reference, note that $2Q - I$ is a unitary matrix, because Q is a projection. All of the matrix elements of Q are equal to $1/d$, so we can write (8.21) explicitly as

$$\alpha_j = \left(\frac{2}{d} - 1\right)\beta_j + \frac{2}{d}\sum_{i \neq j} \beta_i. \tag{8.22}$$

On the full graph Γ, we must deal with the fact that the edges do not have a preferred orientation. To account for this, we decompose each edge into a pair of *bonds*, each carrying one of the possible orientations. For each bond j we assign a coordinate x_j which parametrizes the bond in the direction of its orientation, with $x_j = 0$ at the initial vertex. If Γ has m edges, then this gives coordinates $x_1, \ldots, x_{2m}$, covering each edge twice in opposite directions. In terms of these coordinates, an eigenfunction ψ with eigenvalue $\kappa^2 > 0$ can be written locally as

$$\psi_j(x_j) = \gamma_j e^{i\kappa x_j} \tag{8.23}$$

on the jth bond, where $\gamma_j \in \mathbb{C}$.

To impose the scattering relations at each vertex, we introduce some notation for relationships between bonds:

(i) $\neg j$ denotes the partner bond to j carrying the opposite orientation on the same edge.
(ii) $i \prec j$ means that i precedes j (the initial vertex of j is the final vertex of i).

For each bond j we let ℓ_j denote the length of the corresponding edge.

The term $\gamma_j e^{i\kappa x_j}$ in (8.23) is outgoing from the initial vertex of bond j. In contrast to (8.18), the incoming solutions are parametrized with variables x_i such that $x_i = \ell_i$ at the central vertex. Thus, in the bond notation, the formula (8.22) translates to

$$\gamma_j = \left(\frac{2}{d_j} - 1\right)\gamma_{(\neg j)} e^{i\kappa\ell_j} + \frac{2}{d_j} \sum_{i \prec j, i \neq \neg j} \gamma_i e^{i\kappa\ell_i}, \tag{8.24}$$

where d_j is the degree of the initial vertex of bond j.

To simplify (8.24), we introduce the *bond scattering matrix*

$$S_{ji} = \begin{cases} 2/d_j - 1, & i = \neg j, \\ 2/d_j, & i \prec j \text{ with } i \neq \neg j, \\ 0, & \text{else}, \end{cases}$$

and the diagonal length matrix

$$L_{jk} := \ell_j \delta_{jk}$$

(which is $2m$-dimensional, with each length appearing twice). Then (8.24) translates to

$$\gamma = S e^{i\kappa L} \gamma. \tag{8.25}$$

The matrix S is unitary, by the unitarity of the vertex scattering operator $2Q - I$ from (8.21). Thus $Se^{i\kappa L}$ is unitary as well.

Since $-\Delta$ has an eigenvalue at $\kappa^2 > 0$ if and only if the system (8.25) has a nontrivial solution, we immediately obtain the following:

Theorem 8.13. *The eigenvalues $\lambda = \kappa^2 > 0$ of the Neumann–Kirchoff Laplacian are equal, with multiplicities, to the positive roots of*

$$F(t) := \det(I - Se^{itL}) = 0.$$

The equation $F(t) = 0$ is called the *secular equation* for Γ. To prove that this formula captures the correct multiplicities, we first analyze the eigenvalues of the unitary matrix Se^{itL}. The following lemma is an adaptation of a classical result for

analytic families of normal matrices, specialized to the unitary case. See, e.g., Kato [49, Thm. II.1.10] for a proof.

Lemma 8.14. *Let $U(t)$ be a family of unitary $n \times n$ matrices depending analytically on $t \in \mathbb{R}$. There exists an orthonormal basis of eigenvectors $\{v_j(t)\}$, with corresponding eigenvalues $\{e^{i\theta_j(t)}\}$, such that the functions $v_j : \mathbb{R} \to \mathbb{C}^n$ and $\theta_j : \mathbb{R} \to \mathbb{R}$ are real analytic.*

By Lemma 8.14, we can assume that the eigenvalues of Se^{itL} take the form $e^{i\theta_j(t)}$, with a corresponding orthonormal basis $\{v_j(t)\}$ for $\mathbb{C}^{2m}$, with θ_j and v_j being real analytic functions. Differentiating the equation,

$$Se^{itL} v_j = e^{i\theta_j} v_j, \tag{8.26}$$

with respect to t gives

$$Se^{itL}\Big[iLv_j + v_j'\Big] = e^{i\theta_j}\Big[i\theta_j' v_j + v_j'\Big].$$

Taking the inner product of this equation with (8.26), and using the fact that $\overline{v_j} \cdot v_j' = 0$, because v_j is a unit vector, we obtain

$$\theta_j' = \overline{v_j} \cdot L v_j. \tag{8.27}$$

Proof of Theorem 8.13 It is clear from (8.25) that κ^2 is an eigenvalue if and only if $F(\kappa) = 0$. By the definition of $\theta_j(t)$ given above,

$$F(t) = \prod_{j=1}^{2m} \Big(1 - e^{i\theta_j(t)}\Big). \tag{8.28}$$

The multiplicity of κ^2 as an eigenvalue is the number of values of j for which $\theta_j(\kappa) \in 2\pi\mathbb{Z}$. By (8.27) and the positivity of L, we have $\theta_j'(t) > 0$ for all t. Hence the zeros of $\theta_j(t) = 2\pi k$ are simple for each $k \in \mathbb{Z}$. Therefore, the eigenvalue multiplicity is equal to the degree of vanishing of $F(t)$ at the root. □

Example 8.15. Consider a flower graph with two loops of lengths ℓ_1, ℓ_2, as in Example 8.11. Assign bonds $\{1, 2\}$ to the first loop, and $\{3, 4\}$ to the second. Since all bonds are connected to the central vertex of degree 4, the bond scattering matrix has the form

$$S = \frac{1}{2}\begin{pmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix}.$$

The four eigenvalues of Se^{itL} are given by $e^{it\ell_1}$, $e^{it\ell_2}$, and $\pm e^{it(\ell_1+\ell_2)/2}$, in accordance with Lemma 8.14. The secular equation is thus given by

$$(1 - e^{it\ell_1})(1 - e^{it\ell_2})(1 - e^{it(\ell_1+\ell_2)}) = 0.$$

Reading off the positive roots gives the full spectrum,

$$\sigma(-\Delta) = \{0\} \cup \left(\frac{2\pi\mathbb{N}}{\ell_1}\right)^2 \cup \left(\frac{2\pi\mathbb{N}}{\ell_2}\right)^2 \cup \left(\frac{2\pi\mathbb{N}}{\ell_1+\ell_2}\right)^2.$$

♢

8.5.1 Weyl Law

In Section 6.5, we saw that the asymptotic distribution of eigenvalues of the Dirichlet Laplacian on a bounded open set satisfies an asymptotic power law with exponent equal to half the dimension. In this section we will establish a quantum graph version of the Weyl law. This is a sharp version of the asymptotic

$$\#\{\lambda \in \sigma(-\Delta) \le t\} \sim \frac{\ell(\Gamma)}{\pi} t^{\frac{1}{2}},$$

reflecting the fact that a quantum graph is effectively one dimensional with volume $\ell(\Gamma)$.

Theorem 8.16. *Let Γ be a finite quantum graph with m edges and total length $\ell(\Gamma)$, with $-\Delta$ the Neumann–Kirchoff Laplacian. Assuming that eigenvalues are counted with multiplicity,*

$$\#\big\{\kappa^2 \in \sigma(-\Delta) : a \le \kappa \le b\big\} = \frac{\ell(\Gamma)}{\pi}(b-a) + O(1),$$

uniformly for $0 < a < b$.

Proof Let $\theta_j(t)$ be the phases of the eigenvalues of the unitary matrix Se^{itL}, as defined in the proof of Theorem 8.13. Each eigenvalue κ^2, counted with multiplicity, corresponds to a point where $\theta_j \in 2\pi\mathbb{Z}$ for some particular j. Between a and b, the number of such points for each j is approximately given by $[\theta_j(b) - \theta_j(a)]/2\pi$, with an error of at most 1. Thus,

$$\#\big\{\kappa^2 \in \sigma(-\Delta) : a \le \kappa \le b\big\} = \frac{1}{2\pi}\sum_{j=1}^{2m}\big[\theta_j(b) - \theta_j(a)\big] + R(a,b),$$

with $|R(a,b)| \le 2m$.

Using the derivative formula (8.27), we can write

$$\frac{1}{2\pi}\sum_{j=1}^{2m}\big[\theta_j(b)-\theta_j(a)\big] = \frac{1}{2\pi}\sum_{j=1}^{2m}\int_a^b \theta_j'(t)\,dt$$
$$= \frac{1}{2\pi}\sum_{j=1}^{2m}\int_a^b \overline{v_j(t)}\cdot Lv_j(t)\,dt,$$

where $v_j(t)$ is the jth eigenfunction branch. Since $\{v_j(t)\}$ is an orthonormal basis of $\mathbb{C}^{2m}$ for each t,

$$\sum_{j=1}^{2m}\overline{v_j(t)}\cdot Lv_j(t) = \operatorname{tr} L = 2\ell(\Gamma)$$

independent of t. Therefore,

$$\frac{1}{2\pi}\sum_{j=1}^{2m}\big[\theta_j(b)-\theta_j(a)\big] = \frac{\ell(\Gamma)}{\pi}(b-a),$$

and the result follows. □

8.6 Exercises

8.1. Let Γ be a compact quantum graph with Laplacian $-\Delta$ defined using Neumann–Kirchoff vertex conditions. Assume that Γ is equilateral, meaning that all edges have the same length ℓ. A variant of the combinatorial Laplacian L on Γ is the normalized Laplacian given by

$$\mathcal{L} := D^{-\frac{1}{2}}LD^{-\frac{1}{2}} = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}},$$

where A and D are the adjacency and degree matrices of Γ. For $\sigma \neq \pi\mathbb{Z}/\ell$, prove that $\sigma^2 \in \sigma(-\Delta)$ if and only if $1-\cos(\sigma\ell)$ is an eigenvalue of $\mathcal{L}$.

8.2. Let Γ be a compact quantum graph with Neumann–Kirchoff Laplacian $-\Delta$ and eigenvalues $\{\lambda_n\}$ arranged in increasing order. Let $\{\mu_n\}$ be the eigenvalues of the Laplacian defined by imposing Dirichlet boundary conditions at a single vertex v_0 (leaving all other vertices unchanged).

(a) Prove a min–max formula for μ_n.

(b) Use the min–max formula to show that for all n, the eigenvalues satisfy

$$\lambda_n \le \mu_n \le \lambda_{n+1}.$$

8.3. For a compact quantum graph Γ, let Γ_0 be the corresponding graph with Dirichlet conditions imposed at all vertices (effectively splitting the graph of a disjoint collection of intervals). Using the result of Exercise 8.2 to compare the eigenvalue counting functions of Γ and Γ_0, provide a direct proof of the Weyl asymptotic:

$$\#\{\lambda \in \sigma(\Gamma) : \lambda \le t\} = \frac{\ell(\Gamma)}{\pi}\sqrt{t} + O(1).$$

8.4. Let Γ be a flower graph with lengths $\ell_1, \dots, \ell_m$ where $m \ge 2$. Suppose the Neumann–Kirchoff eigenvalues are arranged in increasing order, starting from $\lambda_1 = 0$. If the total length $\ell = \sum \ell_j$ is fixed, prove that λ_2 is maximized in the equilateral case, where $\ell_j = \ell/m$ for all j.

8.5. Let Γ be a compact quantum graph with vertices $\mathcal{V}$. Given a function $\alpha : \mathcal{V} \to \mathbb{R}$, the analog of Robin boundary conditions are the vertex conditions

$$\sum_{e \ni v} f_e'(v) = \alpha(v) f(v),$$

for all $v \in \mathcal{V}$. Find the quadratic form on $H^1(\Gamma)$ that corresponds to these conditions, and use it to prove that the Robin Laplacian is well defined as a self-adjoint operator on $L^2(\Gamma)$.

8.6. Suppose that Γ and $\tilde{\Gamma}$ are compact quantum graphs, and that Γ is obtained from $\tilde{\Gamma}$ by identifying a pair of vertices, as in Corollary 8.9(a). If $\{\lambda_k\}$ and $\{\tilde{\lambda}_k\}$ are the corresponding eigenvalues in increasing order, prove that

$$\lambda_k \le \tilde{\lambda}_{k+1}$$

for all k.

Notes

For an introduction to combinatorial graph theory, see Chung [22]. Additional background on the spectral theory of quantum graphs may be found in Kuchment [53], Berkolaiko and Kuchment [11], or Berkolaiko [10]. It is possible to define spectral theory in a hybrid context where functions on a metric graph take both discrete values at the vertices and continuous values on the edges. This setting, which is motivated by applications in number theory, is developed in Baker and Rumely [7].

Chapter 9
Spectral Theory on Manifolds



In previous chapters, we have studied the spectral theory of a variety of operators based on the Laplacian on $\mathbb{R}^n$. Many modern applications of spectral theory, in differential geometry, topology, and number theory for example, involve the Laplacian associated with a Riemannian metric. In many contexts, this geometric version of the Laplacian admits a natural self-adjoint extension, and its spectral theory can therefore be used to define geometric invariants.

The goal of this chapter is to lay the groundwork for the study of "spectral geometry," the study of the spectral theory of naturally defined geometric operators. We will give a very basic introduction to Riemannian geometry, explain how the Laplacian is defined, and develop some basic tools for its analysis. This discussion is intended to be self-contained, so its scope is necessarily rather limited. The most serious omission is the topic of curvature. Although curvature is of central importance in spectral geometry, it requires more technical background than our brief introduction will allow. Suggestions for further reading are given in the notes at the end of the chapter.

9.1 Smooth Manifolds

A *topological manifold* is a second countable Hausdorff space which is locally homeomorphic to $\mathbb{R}^n$. For the benefit of readers who do not know this topology background, the essential consequence of this definition is that a manifold admits a set of local coordinate parametrizations modeled on $\mathbb{R}^n$, such that the changes of coordinates between parametrizations are smooth.

In order to avoid a lengthy digression into topological details, we will use the equivalent but more elementary definition of Guillemin and Pollack [39], where manifolds are defined as subsets of Euclidean space. This approach relies on the *induced topology* on a subset $A \subset \mathbb{R}^q$. In the induced topology, a subset of A is

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1_9

open if and only if it is the intersection with A of an open set in $\mathbb{R}^q$. For example, the set A itself is both closed and open in the induced topology on A.

A smooth structure is induced on A in a similar way. A function $f : A \to \mathbb{R}^n$ is smooth if each point $p \in A$ admits an open neighborhood $U \subset \mathbb{R}^q$ on which there exists a smooth function $\tilde{f} : U \to \mathbb{R}^n$ such that the restriction of $\tilde{f}$ to A agrees with f. A *diffeomorphism* between subsets of $\mathbb{R}^q$ and $\mathbb{R}^n$ is an invertible map which is smooth in both directions.

Definition 9.1. A *manifold* of dimension n is a subset $M \subset \mathbb{R}^q$, equipped with the induced topology described above, which is locally diffeomorphic to $\mathbb{R}^n$.

The condition "locally diffeomorphic to $\mathbb{R}^n$" means that each point of M has a neighborhood U and a diffeomorphism ψ from U to an open set $V \in \mathbb{R}^n$. This map ψ, called a *coordinate chart*, is typically written in component form as $\psi = (x^1, \ldots, x^n)$. The inverse map ψ^{-1} defines a parametrization of U using the coordinates $(x^1, \ldots, x^n)$, as illustrated in Figure 9.1. The fact that coordinate indices are raised is part of a useful convention which we will explain below.

A set of coordinate charts covering M is called an *atlas*. If two coordinate charts (ψ_1, U_1) and (ψ_2, U_2) overlap, then the change of coordinates is given by a *transition map*

$$\psi_1 \circ \psi_2^{-1} : \psi_2(U_1 \cap U_2) \to \psi_1(U_1 \cap U_2),$$

which is a diffeomorphism.

The implicit function theorem provides a good source of examples of manifolds. For $n < q$, suppose that $F : \mathbb{R}^q \to \mathbb{R}^{q-n}$ is a smooth function and that dF has maximal rank $q-n$ at all points in the preimage of $a \in \mathbb{R}^{q-n}$. Then $F^{-1}\{a\}$ is locally diffeomorphic to $\mathbb{R}^n$, by the implicit function theorem, and is hence a manifold of dimension n.

Smoothness of maps between manifolds can be formulated in terms of local coordinates. A function $f : M_1 \to M_2$ is smooth if and only all of its realizations in local coordinates are smooth. Two manifolds are considered equivalent if they are related by a global diffeomorphism.

Because each point has a neighborhood diffeomorphic to an open subset of $\mathbb{R}^n$, a manifold has no boundary points. To allow for a boundary, we modify the definition

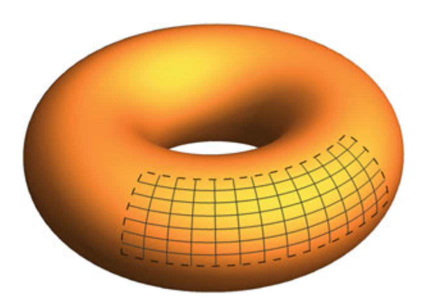


Fig. 9.1 A coordinate chart on a surface in $\mathbb{R}^3$

by taking the closed half-space $\overline{\mathbb{R}^n_+} := \{x^n \geq 0\} \subset \mathbb{R}^n$ as a model for the local neighborhoods.

Definition 9.2. A *manifold with boundary* of dimension n is a closed subset $\overline{\Omega} \subset \mathbb{R}^q$ that is locally diffeomorphic to $\overline{\mathbb{R}^n_+}$.

In standard usage, term "manifold" refers only to the case of Definition 9.1, without boundary. However, the term "manifold with boundary" is inclusive, i.e., the boundary is allowed to be empty as a special case. This is consistent with Definition 9.2, because the neighborhoods of $\overline{\mathbb{R}^n_+}$ could all lie in the interior.

For a manifold with boundary $\overline{\Omega}$, the interior Ω is a manifold in the ordinary sense. Points in the boundary $\partial\Omega$ are necessarily mapped to the hyperplane $\partial\mathbb{R}^n_+ = \{x^n = 0\}$ in local coordinate patches. It follows that the restrictions of coordinate charts to $\partial\mathbb{R}^n_+$ form an atlas for $\partial\Omega$, giving the boundary the structure of a manifold of dimension $n-1$.

By definition, a smooth map into a neighborhood of $\overline{\mathbb{R}^n_+}$ admits (locally at least) a smooth extension to an open subset of $\mathbb{R}^n$. Thus a manifold with boundary admits a realization as a closed subset of a larger manifold without boundary.

9.1.1 Tangent and Cotangent Vectors

In vector calculus, the tangent vectors to a manifold M embedded in $\mathbb{R}^q$ would be defined in terms of derivatives of parametrized curves. That is, a tangent vector to M at the point p is the derivative $\dot{\gamma}(t_0) \subset \mathbb{R}^q$ of a smooth parametrized curve $\gamma : (a,b) \to M$, for which $\gamma(t_0) = p$.

The problem with this definition is that it depends explicitly on the embedding. To avoid this dependence, we focus on the fact that a curve also defines a directional derivative. With M and γ as above, for $f \in C^\infty(M)$ the derivative of $f \circ \gamma$ at t_0 depends on the curve only through $\gamma(t_0)$ and $\dot{\gamma}(t_0)$, by the chain rule. We can thus identify $\dot{\gamma}(t_0)$ with the directional derivative map,

$$f \mapsto \frac{d}{dt}(f \circ \gamma)(t_0). \tag{9.1}$$

This interpretation is the motivation for the following:

Definition 9.3. A *tangent vector* at a point $p \in M$ is a linear map

$$v : f \in C^\infty(M) \mapsto v(f) \in \mathbb{R},$$

with the following properties. For $f, g \in C^\infty(M)$,

(i) $v(f) = v(g)$ provided $f = g$ in some neighborhood of p (*locality*);
(ii) $v(fg) = f(p)v(g) + g(p)v(f)$ (*derivation*).

Under Definition 9.3, we can still interpret tangent vectors to curves via (9.1), and we continue to use the notation $\dot{\gamma}(t_0)$ for such "velocity" vectors.

Directional derivatives in calculus are often decomposed as linear combinations of partial derivatives. The analog of a partial derivative on a manifold is a derivative with respect to local coordinates. With the components of a coordinate chart $\psi : U \subset M \to V \subset \mathbb{R}^n$ written as $(x^1, \ldots, x^n)$, we define the tangent vectors

$$\partial_j f := \frac{\partial}{\partial x_j}\big(f \circ \psi^{-1}\big),$$

for $j = 1, \ldots, n$. It follows from the fact that ψ is a diffeomorphism that $\{\partial_j\}|_p$ forms a basis for T_pM when $p \in U$. The collection $\{\partial_j\}$ is called a *coordinate frame*. In Figure 9.1, we can visualize this frame as a set of tangent vectors running parallel to the coordinate lines.

The space of tangent vectors to M at p is denoted by T_pM, and the union of tangent spaces is the *tangent bundle*,

$$TM := \bigcup_{p \in M} T_pM.$$

The coordinate frames give a natural set of coordinate patches for TM, so the tangent bundle is also a manifold.

Notice that the indices on the frame vectors ∂_j are lowered, while coordinate indices are raised. The convention, introduced by Albert Einstein, is that pairs of upper/lower indices on the same side of an equation are implicitly summed over:

$$a^i b_i := \sum_{i=1}^{n} a^i b_i.$$

We will adopt this Einstein summation convention for the remainder of the chapter.

A *vector field* on M is a smooth map $v : M \to TM$ such that $v_p \in T_pM$ for each $p \in M$. With respect to a local coordinate frame, the vector field has the form

$$v = v^i \partial_i,$$

for some real-valued functions v^i defined on the coordinate patch. The smoothness of the vector field is equivalent to the condition that the coefficients v^i are smooth in each coordinate patch. Since tangent vectors are identified with directional derivatives, a vector field can be interpreted as a first-order differential operator on $C^\infty(M)$.

The vector space dual of T_pM is denoted by T_p^*M, and its elements are called *cotangent vectors*. The union of all of the cotangent spaces forms the *cotangent bundle*

$$T^*M := \bigcup_{p\in M} T_p^*M.$$

A natural way to obtain cotangent vectors is by differentiation of functions. Given a smooth function $f \in C^\infty(M;\mathbb{R})$, the differential $df_p \in T_p^*M$ is the evaluation map

$$df_p(v) := v(f), \tag{9.2}$$

for $v \in T_pM$. Applying this definition to a set of local coordinate functions $(x^1,\ldots,x^n)$ gives a local coordinate frame $(dx^1,\ldots,dx^n)$ for T^*M. The coordinate frames for TM and T^*M are dual bases, in the sense that

$$dx^i(\partial_j) = \delta^i_j, \tag{9.3}$$

where δ^i_j is the usual Kronecker delta, adapted to the Einstein index convention. By (9.3), the pairing of a cotangent vector $\xi = a_j dx^j$ with the tangent vector $v = v^i\partial_i$ takes the very natural form,

$$\xi.v = a_j v^j.$$

The summation convention also makes it easy to apply the chain rule by simply matching index pairs. For example, for two sets of coordinates $(x^1,\ldots,x^n)$ and $(y^1,\ldots,y^n)$, the coordinate frames are related by

$$dy^j = \frac{\partial y^j}{\partial x^i}dx^i, \qquad \partial_{y^j} = \frac{\partial x^i}{\partial y^j}\partial_{x^i}.$$

The above definitions of tangent and cotangent vectors apply also to a manifold with boundary $\overline{\Omega}$. If a boundary neighborhood is parametrized by coordinates $(x^1,\ldots,x^n)$, with $x_n \geq 0$, then this defines a local coordinate frame $\{\partial_j\}$ for $T\overline{\Omega}$. At a boundary point $p \in \partial\Omega$ the tangent space is still isomorphic to $\mathbb{R}^n$, but we can distinguish the vectors as inward, tangent to $\partial\Omega$, or outward, according to the sign of the coefficient of ∂_n.

9.1.2 Partition of Unity

For many of the arguments given in this chapter, we will reduce calculations to local computations in $\mathbb{R}^n$ using coordinate patches. To make this work, we need a way to piece together objects defined in local coordinates into a structure defined on the full manifold. In this section we will describe the essential tool for this purpose, a smooth partition of unity.

By the Heine–Borel property, closed and bounded subsets of $\mathbb{R}^n$ are compact. This implies that a manifold M defined according to Definition 9.1 is *locally compact*, meaning that each point has a neighborhood whose closure is compact. Any open cover of M can be reduced to a locally finite cover, meaning that each point is contained in only a finite number of sets in the cover. In particular, we can assume that the coordinate atlas on M is locally finite.

Lemma 9.4 (Partition of Unity). *Let $\{U_j, \psi_j\}$ be a locally finite atlas for M. There exists a collection of functions $\chi_j \in C_0^\infty(M)$ such that χ_j has support in U_j and*

$$\sum_j \chi_j = 1.$$

Proof For each j, the set

$$F_j := \left\{x \in U_j : x \notin U_k \text{ for all } k \neq j\right\}$$

is a closed subset of U_j. We can choose a positive function $h_j \in C_0^\infty(M)$ with support in U_j such that $h_j = 1$ on F_j; this guarantees that at least one h_j is nonzero at each point of M. We then set

$$\chi_j(x) := \frac{h_j(x)}{\sum_j h_j(x)}.$$

There is no convergence issue for the sum in the denominator, because only finitely many terms are nonzero for each x. □

9.2 Riemannian Metrics

A *Riemannian metric* is a family of inner products $g(\cdot, \cdot)$ on T_pM that varies smoothly as a function of p. In a coordinate frame, the metric is represented by a symmetric, positive matrix with elements

$$g_{ij} := g(\partial_i, \partial_j).$$

For example, in Cartesian coordinates, the Euclidean norm is given by the identity matrix, $g_{ij} = \delta_{ij}$.

The condition that g is smooth means that the components g_{ij} are smooth functions in each coordinate patch. In terms of the coordinate basis for T_p^*M, we can use the summation convention to write

$$g = g_{ij}\, dx^i \otimes dx^j.$$

The norm on T_pM associated with g will be denoted by

$$|v|_g := \sqrt{g(v, v)}$$

for $v \in T_pM$.

Given a Riemannian metric, we can define the speed of a curve $\gamma : I \to M$ as $|\dot{\gamma}|_g$. Assuming that I is a finite interval, the arclength is then given by the usual formula,

$$\ell(\gamma) := \int_I |\dot{\gamma}|_g \, dt. \tag{9.4}$$

Let s denote the displacement along the curve, measured from a starting time $t_0 \in I$,

$$s(t) := \int_{t_0}^{t} |\dot{\gamma}|_g \, dt'. \tag{9.5}$$

In local coordinates, this formula reduces to

$$s(t) := \int_{t_0}^{t} \sqrt{g_{ij}\dot{\gamma}^i\dot{\gamma}^j} \, dt'.$$

We often write this relationship between metric and length in abbreviated form as an expression for the infinitesimal length element squared,

$$ds^2 = g_{ij} dx^i dx^j. \tag{9.6}$$

This gives a convenient shorthand for specifying the metric in coordinate form. For example, the Euclidean metric in $\mathbb{R}^2$ is written as $ds^2 = dx^2 + dy^2$ in Cartesian coordinates, or $ds^2 = dr^2 + r^2 d\theta^2$ in polar.

Example 9.5. The historical prototype for a Riemannian manifold is a regular surface embedded in $\mathbb{R}^3$, with metric induced by the three-dimensional dot product. Here are a few cases of induced surface metrics:

1. For a graph $z = f(x, y)$, we use the obvious parametrization,

$$\sigma(x, y) = (x, y, f(x, y)).$$

The tangent vectors ∂_x and ∂_y correspond to three-dimensional vectors

$$\sigma_x = (1, 0, f_x), \qquad \sigma_y = (0, 1, f_y),$$

respectively, where the subscripts denote partial derivatives. The components of the induced metric are given by

$$g_{11} = \sigma_x \cdot \sigma_x, \quad g_{12} = g_{21} = \sigma_x \cdot \sigma_y, \quad g_{22} = \sigma_y \cdot \sigma_y.$$

In the notation (9.6) the metric would be written as

$$ds^2 = \left(1 + f_x^2\right)dx^2 + 2f_x f_y\, dx\, dy + \left(1 + f_y^2\right)dy^2.$$

2. Consider a surface of revolution defined by $z = h(r)$, with the polar coordinate parametrization,

$$\sigma(r, \theta) = (r\cos\theta, r\sin\theta, h(r)).$$

The tangent vectors ∂_r and ∂_θ correspond to the three-dimensional vectors

$$\sigma_r = (\cos\theta, \sin\theta, h_r), \qquad \sigma_\theta = (-r\sin\theta, r\cos\theta, 0).$$

Taking dot products as above gives the induced metric

$$ds^2 = \left(1 + h_r^2\right)dr^2 + r^2\, d\theta^2.$$

3. The standard parametrization of the unit sphere $\mathbb{S}^2$ is

$$\sigma(\varphi, \theta) := (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi),$$

where φ is the polar angle and θ the azimuth. Taking coordinate derivatives and dot products gives the metric

$$ds^2 = d\varphi^2 + \sin^2\varphi\, d\theta^2.$$

$\diamondsuit$

The definition (9.4) of arclength determines a metric structure on M, with distance function

$$\operatorname{dist}(p, q) := \inf\left\{\ell(\gamma) :\ \gamma \in C^\infty([0, 1], M), \gamma(0) = p, \gamma(1) = q\right\}. \tag{9.7}$$

We will see below that the metric topology generated by $\operatorname{dist}(\cdot, \cdot)$ is compatible with the preexisting manifold topology on M. By (9.6), we can interpret the Riemannian metric as the infinitesimal form of the distance function $\operatorname{dist}(\cdot, \cdot)$.

To define Riemannian metrics on manifolds with boundary, we impose the smoothness requirement up to the boundary, meaning that the matrix elements g_{ij} admit local smooth extensions across the boundary. Under this definition, a Riemannian manifold with boundary can be realized as a closed subset (with smooth boundary) within an ordinary Riemannian manifold.

The restriction of g to vectors which are tangent to the boundary defines a metric h on ∂M. In local coordinates, (with the boundary given by $\{x_n = 0\}$) this boundary metric is simply given by

$$h_{ij} = g_{ij}\big|_{x_n=0},$$

for $i, j \in \{1, \ldots, n-1\}$.

9.2.1 Geodesics and the Exponential Map

To understand the distance structure defined in (9.7), it is helpful to work out the necessary local conditions for a path to minimize distance. For this purpose we use the Euler–Lagrange equations for a critical point of the length functional such as $\ell(\cdot)$. These variational equations can be worked out explicitly by considering families of perturbations of a curve γ. It is possible (and generally preferable) to do this in a coordinate-free way, through the introduction of connections and covariant derivatives. However, since we will not need such technology elsewhere in this chapter, we will simplify the discussion by developing the variational equations in local coordinates.

Suppose $\gamma : [0, b] \to M$ is a smooth curve contained in a single coordinate patch, with components $(\gamma^1, \dots, \gamma^n)$ We will assume that γ is parametrized at constant speed c, which by (9.6) means that

$$g_{ij}\dot{\gamma}^i\dot{\gamma}^j = c^2. \tag{9.8}$$

Consider a variation of γ defined as

$$\gamma_h^j := \gamma^j + h\eta^j,$$

where each η^j is a smooth function $[0, b] \to \mathbb{R}$ with $\eta^j(0) = \eta^j(b) = 0$, and $h \in \mathbb{R}$ is sufficiently small, so that γ_h stays within the patch. The length of this perturbed curve is

$$\ell(\gamma_h) = \int_0^b \sqrt{g_{ij}\dot{\gamma}_h^i\dot{\gamma}_h^j}\, dt.$$

The assumption that $\ell(\gamma)$ is minimal implies in particular that $h = 0$ is a critical point of $h \mapsto \ell(\gamma_h)$. Thus gives the variational equation,

$$\frac{d}{dh}\ell(\gamma_h)\Big|_{h=0} = 0, \tag{9.9}$$

for all η.

To compute the derivative in (9.9), we must keep in mind that the coefficients g_{ij} appearing in the integrand also vary along γ_h, so that

$$\frac{d}{dh}g_{ij}\Big|_{h=0} = (\partial_k g_{ij})\eta^k.$$

Under the constant speed assumption (9.8), we thus calculate

$$\frac{d}{dh}\ell(\gamma_h)\Big|_{h=0} = \frac{1}{c}\int_0^b \left[g_{ij}\dot{\gamma}^i\dot{\eta}^j + \frac{1}{2}(\partial_k g_{ij})\eta^k\dot{\gamma}^i\dot{\gamma}^j \right] dt$$

Applying integration by parts to the $\dot{\eta}$ term then gives

$$\frac{d}{dh}\ell(\gamma_h)\Big|_{h=0} = \frac{1}{c}\int_0^b \left[-\frac{d}{dt}\left(g_{ik}\dot{\gamma}^i\right) + \frac{1}{2}(\partial_k g_{ij})\dot{\gamma}^i\dot{\gamma}^j\right]\eta^k\,dt,$$

since η vanishes at the endpoints. It is now clear that (9.9) will be satisfied if and only if

$$\frac{d}{dt}(g_{ik}\dot{\gamma}^i) = \frac{1}{2}\big(\partial_k g_{ij}\big)\dot{\gamma}^i\dot{\gamma}^j. \tag{9.10}$$

To write (9.10) in a more standard form, we take the derivative on the left and multiply by the inverse matrix g^{il}. This yields the *geodesic equation*,

$$\ddot{\gamma}^i + \Gamma^i_{jk}\dot{\gamma}^j\dot{\gamma}^k = 0, \tag{9.11}$$

where Γ^i_{jk} is the *Christoffel symbol*

$$\Gamma^i_{jk} := \frac{1}{2}g^{il}\big[\partial_j g_{lk} + \partial_k g_{lj} - \partial_l g_{jk}\big]. \tag{9.12}$$

It is straightforward to check that any solution of (9.11) satisfies the speed equation (9.8) for some constant c.

Since the length functional does not depend on the choice of coordinates, it follows that the geodesic equation is invariant under changes of coordinates. (The left-hand side of (9.11) is the coordinate representation of the "covariant acceleration" vector.)

Definition 9.6. A *geodesic* on a Riemannian manifold is a curve which satisfies (9.11) in each local coordinate patch.

Example 9.7. Consider the sphere $\mathbb{S}^2$, parametrized as in Example 9.5 with the spherical coordinates ordered as (φ, θ). The only component of the metric that has a nonzero derivative is

$$\partial_1 g_{22} = 2\sin\varphi\cos\varphi.$$

Therefore

$$\Gamma^1_{22} = -\sin\varphi\cos\varphi, \qquad \Gamma^2_{12} = \Gamma^2_{21} = \cot\varphi,$$

and all other components are zero. The geodesic equations (9.11) for a curve $\gamma(t) = (\varphi(t), \theta(t))$ reduce to

$$\ddot{\varphi} = \dot{\theta}^2\sin\varphi\cos\varphi, \qquad \ddot{\theta} = -2\dot{\varphi}\dot{\theta}\cot\varphi.$$

An obvious family of solutions is given by taking $\theta = \theta_0$ (constant) and $\ddot{\varphi} = 0$. This yields a meridian at longitude θ_0. The equator $\varphi = \pi/2$ is the only other case that has a simple expression in this coordinate system, but the symmetry of the sphere implies that all other solutions are also "great circle" routes. ◊

To establish the existence of solutions to the geodesic equation, let us consider $v^i := \dot{\gamma}^i$ and γ^i as dependent variables, so that we can write (9.11) as a first-order system in $2n$ variables,

$$\frac{d}{dt}\begin{Bmatrix} \gamma^i \\ v^i \end{Bmatrix} = \begin{Bmatrix} v^i \\ \Gamma^i_{jk} v^j v^k \end{Bmatrix}. \tag{9.13}$$

The classical existence and uniqueness result for first-order systems is the Picard–Lindelöf theorem (also called Cauchy–Lipschitz). It can be proven as an application the contraction mapping principle, under a range of regularity assumptions. For our purposes, the relevant version is the following, adapted from Lee [58, Thm 7.9]:

Theorem 9.8 (Picard–Lindelöf). *Let U be a domain in $\mathbb{R}^n$, and $f : U \to \mathbb{R}^n$ a smooth map. Given $q_0 \in U$, choose r_0 so that $\overline{B(q_0; 2r_0)} \subset U$. Then there exists $\delta > 0$, depending only on r_0 and the C^1 norm of the restriction of f to $\overline{B(q_0; 2r_0)}$, such that for each $q \in B(q_0, r_0)$, the system of ODE*

$$\frac{dy}{dt} = f(y), \qquad y(0) = q, \tag{9.14}$$

has a unique solution $y_q(t)$ for $|t| < \delta$. The function $(t, q) \mapsto y_q(t)$ is smooth on $(-\delta, \delta) \times B(q_0, r_0)$.

The geodesic system (9.13) is of the form (9.14), with

$$f(x^i, v^i) = (v^i, \Gamma^i_{jk} v^j v^k).$$

The functions Γ^i_{jk} are smooth, so f is clearly bounded with respect to the C^1 norm for x contained within a coordinate neighborhood and $|v|_g$ bounded. The existence and uniqueness result of Theorem 9.8 thus yields the following:

Corollary 9.9. *Given a point $p \in M$, there exists a neighborhood $U \subset T_pM$, containing 0, and $\delta > 0$ such that for each $v \in U$ there is a unique geodesic $\gamma_v : (-\delta, \delta) \to M$ with*

$$\gamma_v(0) = p, \qquad \dot{\gamma}_v(0) = v.$$

The solutions provided by Corollary 9.9 satisfy a useful homogeneity property with regard to the velocity. To see this, suppose that γ_v is a geodesic defined for $|t| < \delta$. For $\alpha > 0$ the rescaled curve $t \mapsto \gamma_v(\alpha t)$ is also a solution of the geodesic equations. By the chain rule,

$$\frac{d}{dt}\gamma_v(\alpha t)\Big|_{t=0} = \alpha v.$$

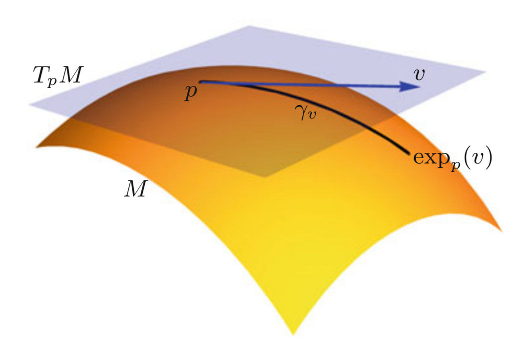


Fig. 9.2 A tangent vector and its image under the exponential map

Therefore, by uniqueness of solutions, $\gamma_{\alpha v}$ exists for $|t| < \delta/\alpha$ and satisfies

$$\gamma_{\alpha v}(t) = \gamma_v(\alpha t). \tag{9.15}$$

From (9.15), we can see that there exists a sufficiently small neighborhood $U \subset T_pM$ in Corollary 9.9 such that $\gamma_v(t)$ exists for all $0 \le t \le 1$ and $v \in U$. On such a neighborhood, we define the *exponential map* $\exp_p : U \to M$ by

$$\exp_p(v) := \gamma_v(1), \tag{9.16}$$

as illustrated in Figure 9.2.

The smooth dependence on initial data from Theorem 9.8 implies that $\exp_p$ is smooth.

From the exponential map we can derive a very useful set of local coordinates on M. For this purpose, we need first to establish local invertibility.

Lemma 9.10. *For each $p \in M$, there is a neighborhood $U_p \subset T_pM$, containing zero, such that the restriction of* $\exp_p$ *to U_p is a diffeomorphism onto its image in M.*

Proof Let us compute the differential of $\exp_p$ at 0. Since T_pM is a vector space, we can naturally identify $T_0(T_pM)$ with T_pM. Under this identification, we can use (9.15) to compute

$$\begin{aligned} d(\exp_p)_0 w &:= \frac{d}{dt} \exp_p(tw)\Big|_{t=0} \\ &= \frac{d}{dt}\gamma_{tw}(1)\Big|_{t=0} \\ &= \frac{d}{dt}\gamma_w(t)\Big|_{t=0} \\ &= w. \end{aligned}$$

Thus, as a map $T_pM \to T_pM$,

$$d(\exp_p)_0 = I. \tag{9.17}$$

Since $d(\exp_p)_0$ is nonsingular, the inverse function theorem implies that the restriction of $\exp_p$ to some neighborhood of 0 is a diffeomorphism. □

To obtain coordinates from the exponential map, we fix an orthonormal basis $\{e_i\}$ for T_pM and parametrize the neighborhood U_p produced in Lemma 9.10 by identifying $(x^1, \ldots, x^n) \in \mathbb{R}^n$ with $x^i e_i \in T_pM$. Composing this with $\exp_p$ yields a *geodesic normal coordinate* patch

$$\psi(x^1, \ldots, x^n) := \exp_p(x^i e_i). \tag{9.18}$$

An example is shown in Figure 9.3.

Within a geodesic normal patch, we define the radial coordinate,

$$r := \sqrt{(x^1)^2 + \cdots + (x^n)^2}. \tag{9.19}$$

The maximal value of r at the point p is called the *injectivity radius*,

$$\operatorname{inj}(p) := \sup\big\{r_0 > 0 : \ \exp_p \ \text{is a diffeomorphism for } r < r_0\big\}. \tag{9.20}$$

For example, the unit sphere pictured in Figure 9.3 has injectivity radius equal to π at every point.

One of the key features of geodesic normal coordinates is the fact that the metric approximates the Euclidean model as closely as possible at the base point.

Lemma 9.11. *In geodesic normal coordinates $(x^1, \ldots, x^n)$, the partial derivatives $\partial_k g_{ij}$ and Christoffel symbols Γ^k_{ij} vanish at $x = 0$, and the metric satisfies*

$$g_{ij} = \delta_{ij} + O(|x|^2). \tag{9.21}$$

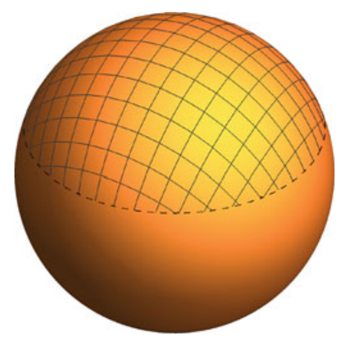


Fig. 9.3 A geodesic normal coordinate patch centered at the north pole of a sphere

Proof By (9.18) and (9.17),

$$\partial_j|_{x=0} = e_j.$$

Since $\{e_j\}$ is orthonormal, this implies that

$$g_{ij}|_{x=0} = \delta_{ij}.$$

For $v = v^j e_j \in U_p$, consider the geodesic $\gamma_v(t)$. In normal coordinates, this is a radial path of the form

$$\gamma_v(t) = (v^1 t, \dots, v^n t). \tag{9.22}$$

The local geodesic equation (9.11) thus implies that

$$\Gamma^i_{jk}\big|_{\gamma_v(t)} v^j v^k = 0,$$

for $i = 1, \dots, n$. This condition applies at $x = 0$ for all choices of v, implying that

$$\Gamma^i_{jk}\big|_{x=0} = 0.$$

It follows that the partial derivatives also vanish at $x = 0$, since

$$\partial_i g_{jk} = g_{jl}\Gamma^l_{ik} + g_{kl}\Gamma^l_{ij}.$$

□

By the definition (9.16) of the exponential map, the geodesic radial coordinate r of a point q corresponds to the length of the radial geodesic from p (the base point) to q. We will show below that the radial path is in fact the shortest route from p to q. First, we need to establish a fundamental property of radial paths. A *geodesic sphere* is defined as a set of the form $\{r = a\}$ within a geodesic normal patch, i.e., for a less than the injectivity radius.

Lemma 9.12 (Gauss). *For $p \in M$, let $\psi : U_p \to M$ be a geodesic normal coordinate patch centered at p. If r denotes the radial coordinate* (9.19) *then ∂_r is the outward unit normal to the geodesic sphere $\{r = a\}$ for each $0 < a < \operatorname{inj}(p)$.*

Proof Suppose that q is a point on the sphere $\{r = a\}$. This means that $q = \exp_p(av)$ for some unit vector $v \in \mathbb{R}^n$. By a change of basis, we can assume that v corresponds to the first basis vector e_1 used to define the coordinate system. This means that $q = (a, 0, \dots, 0)$ in coordinate form, and $\partial_r = \partial_1$ at q. This is the velocity vector of the coordinate curve $\gamma_{e_1}(t) = (t, 0, \dots, 0)$, so ∂_r is a unit vector because $|e_1|_g = 1$.

The set of tangent vectors to $\{r = a\}$ at q is spanned by the coordinate vectors ∂_j with $j = 2, \dots, n$. Thus, to prove that ∂_r is normal to $\{r = a\}$, it suffices to show

that $g_{1j}|_q = 0$ for $j \geq 2$. From the fact that $\gamma_{e_1}(t)$ satisfies the geodesic equation, as written in the form (9.10), we can deduce that

$$\frac{d}{dt} g_{1j} = \frac{1}{2} \partial_j g_{11}.$$

Since $g_{11} = \|\dot{\gamma}_{e_1}\|^2 = 1$, this shows that $\partial_1 g_{1j} = 0$ along γ_{e_1}. By Lemma 9.11, $g_{1j} = 0$ at $x = 0$ for $j \geq 2$. Hence $g_{1j}|_q = 0$. □

Suppose we switch from geodesic normal to the corresponding polar coordinates, by setting $x = r\omega$ with $\omega \in \mathbb{S}^{n-1}$. Lemma 9.12 implies that the metric decomposes as

$$g = dr^2 + h, \tag{9.23}$$

where $h = h(r)$ denotes a family of metrics on $\mathbb{S}^{n-1}$, depending on r. This coordinate system is valid within the geodesic normal patch, and in particular for $r < \operatorname{inj}(p)$.

Example 9.13. Let p be the north pole on the unit sphere $\mathbb{S}^2$. The geodesics originating from p are the meridians, indexed by the azimuthal angle θ. In the case the geodesic polar coordinates are given by

$$(r, \theta) \mapsto (\sin r \cos\theta, \sin r \sin\theta, \cos r),$$

i.e., standard spherical coordinates. (Here r denotes the distance from the north pole, and not the distance from the origin as in the usual spherical coordinates.) The form of the metric was worked out in Example 9.5,

$$ds^2 = dr^2 + \sin^2 r \, d\theta^2. \tag{9.24}$$

In this case h is a family of metrics on the circle given by $h(r) = \sin^2 r \, d\theta^2$. By symmetry, geodesic normal coordinates centered at any point of the sphere have the form (9.24). ♢

Example 9.14. The *Poincaré disk* $\mathbb{B}$ is a manifold based on the unit disk in $\mathbb{R}^2$, equipped with the *hyperbolic* metric. In polar coordinates (ρ, θ) on $\mathbb{B}$, this has the form

$$ds^2 := \frac{4}{(1-\rho^2)^2}\big(d\rho^2 + \rho^2 \, d\theta^2\big). \tag{9.25}$$

Since the metric is invariant under rotation, we would expect geodesics starting at the origin to be radial.

To verify this, first note that a radial curve $t \mapsto (\rho(t), \theta_0)$ has unit speed if

$$\frac{2\dot{\rho}}{1-\rho^2} = 1.$$

The unique solution with $\rho(0) = 0$ is $\rho(t) = \tanh(t/2)$. The radial coordinate r is thus related to ρ by

$$\rho = \tanh(r/2).$$

Substituting this into (9.25) gives

$$ds^2 = dr^2 + \sinh^2 r \, d\theta^2. \tag{9.26}$$

In the (r, θ) coordinates, we have $\Gamma^1_{11} = \Gamma^2_{11} = 0$. It follows that radial curves satisfy the geodesic equation (9.11).

If $\mathbb{B}$ is identified with the unit disk in $\mathbb{C}$, then the Möbius transformations that map $\mathbb{B}$ to itself are isometries of the hyperbolic metric. This group of symmetries implies that the hyperbolic metric has the local form (9.26) at any point. ♢

We are now finally prepared for the main result of this section, which says that sufficiently short geodesics are length-minimizing curves.

Theorem 9.15. *For $p \in M$ and $v \in T_pM$ such that $|v|_g < \operatorname{inj}(p)$, the radial geodesic from p to $q := \exp_p(v)$ is length-minimizing. Conversely, the minimal length curves originating from p are radial within the geodesic polar neighborhood.*

Proof Suppose that $|v|_g = a$, so that the radial geodesic $\gamma_v(t)$ connecting p to q has length a. Let $\gamma : [0, 1] \to U_p$ be another curve connecting p to q. If the image of γ lies within the region $\{r < \operatorname{inj}(p)\}$, then the curve can be written in polar coordinates as $\gamma(t) = (r(t), \omega(t))$. By (9.23),

$$\begin{aligned} |\dot\gamma|_g^2 &= \dot r^2 + h(\dot\omega, \dot\omega) \\ &\geq \dot r^2. \end{aligned}$$

Thus,

$$\begin{aligned} \ell(\gamma) &= \int_0^1 |\dot r| \, dt \\ &\geq \int_0^1 \dot r \, dt \\ &= a. \end{aligned}$$

On the other hand, if γ passes into the region $\{r > a\}$ (and possibly out of the geodesic polar neighborhood), then the same argument shows that $\ell(\gamma) > a$. Therefore, the radial path has the shortest length among all possible curves connecting p to q, and hence $\operatorname{dist}(p, q) = a$. □

9.2.2 Completeness

Theorem 9.15 shows that if r is the radial coordinate defined within a geodesic normal patch centered at p, then $\mathrm{dist}(p,\cdot) = r$ for $r < \mathrm{inj}(p)$. This implies in particular that metric balls $B(p,r)$ are open in the topology of M, and that each open neighborhood of p contains arbitrarily small metric balls. Thus, the metric topology on M defined by (9.7) is therefore compatible with the original manifold topology.

We say that a Riemannian manifold M is *complete* if it is complete as a metric space. Complete manifolds with boundary are defined by the same condition.

There is an alternative formulation of completeness in terms of geodesics. A Riemannian manifold is *geodesically complete* if all geodesics are defined for all $t \in \mathbb{R}$. The link between these notions is a fundamental result in Riemannian geometry.

Theorem 9.16 (Hopf–Rinow). *For a Riemannian manifold M, the following are equivalent:*

(a) *M is complete.*
(b) *M is geodesically complete.*
(c) *At some $p \in M$, $\exp_p$ is defined on all of T_pM.*
(d) *Closed and bounded subsets of M are compact (the Heine–Borel property).*

The proof is slightly too technical for us to include here, but is covered in all standard differential geometry texts. See, e.g., Petersen [66, §5.7.1] for a version that includes the Heine–Borel property, which is sometimes omitted from the statement. The Heine–Borel criterion for completeness can be easily extended to the case of a manifold with boundary. Properties (b) and (c) obviously do not hold unless the boundary is empty.

Example 9.17. For the *Poincaré disk* $\mathbb{B}$ introduced in Example 9.14, we saw that unit speed geodesics starting at $\rho = 0$ have the form $\rho(t) = \tanh(t/2)$, $\theta(t) = \theta_0$. Since this solution exists for all $t \in [0,\infty)$, the exponential map at the origin is defined for all tangent vectors. Therefore, $\mathbb{B}$ is complete by property (c) of Theorem 9.16.

Using the Möbius symmetry, and the fact that

$$|z| = \tanh(\mathrm{dist}(0,z)/2),$$

we can derive an explicit formula for the distance function for the Poincaré metric,

$$\mathrm{dist}(w,z) = \log\left[\frac{|w-z| + \sqrt{|w|^2|z|^2 - 2w\cdot z + 1}}{\sqrt{(1-|w|)^2(1-|w|^2)}}\right].$$

From this expression, it is possible to verify directly that $\mathbb{B}$ is complete as a metric space. ♢

In addition to Hopf–Rinow, we need to cite one more technical result without proof. Many of the arguments in later sections will make use of geodesic coordinates, and it will be necessary to have some control over the size of normal neighborhoods.

Lemma 9.18. *For a complete Riemannian manifold M, the injectivity radius* $\mathrm{inj} : M \to \mathbb{R}_+$ *is continuous. In particular, on a compact set the injectivity is bounded below by a positive constant.*

The proof is not difficult, but it does require some additional machinery. For details see Klingenberg [52, Prop. 2.1.10].

9.3 The Laplacian

In $\mathbb{R}^n$ the Laplacian operator can be defined as the divergence of the gradient of a function. To make a corresponding definition on a Riemannian manifold M, we need to develop the notions of gradient and volume density.

For a function $f \in C^\infty(M)$, the differential $df : M \to T^*M$ is defined by (9.2), independently of the metric. In local coordinates,

$$df = \partial_j f \, dx^j.$$

The metric g provides a natural identification of T_p^*M with T_pM, allowing us to define a corresponding *gradient* map $\nabla f : M \to TM$, such that

$$g(\nabla f(p), v) = df_p(v) = v(f), \tag{9.27}$$

for $v \in T_pM$.

In terms of the local coordinate frame, we can expand $\nabla f = (\nabla f)^j \partial_j$. Setting $v = \partial_k$ in (9.27) then gives

$$g_{kj}(\nabla f)^j = \partial_k f. \tag{9.28}$$

By convention, the inverse matrix to g_{ij} is written with upper indices, so that

$$g^{ik} g_{kj} = \delta^i_j.$$

Applying g^{ik} to (9.28) yields the local coordinate formula for the gradient,

$$\nabla f = g^{jk} \partial_k f \, \partial_j.$$

Similarly, the inner product of the gradients of two functions is given locally by

$$g(\nabla u, \nabla v) = g^{ij} \partial_i u \, \partial_j v.$$

Example 9.19. In polar coordinates (r, θ) for $\mathbb{R}^2$, $ds^2 = dr^2 + r^2\, d\theta^2$. The differential of a function $f(r, \theta)$ is

$$df = f_r\, dr + f_\theta\, d\theta.$$

Since $g^{11} = 1$, $g^{12} = g^{21} = 0$, and $g^{22} = r^{-2}$, the gradient is

$$\nabla f = f_r\, \partial_r + r^{-2} f_\theta\, \partial_\theta.$$

◊

The second ingredient that we need to define the Laplacian is volume density associated with a Riemannian metric. Under a change of local coordinates from $(x^1, \dots, x^n)$ to $(y^1, \dots, y^n)$, the Jacobian formula gives

$$dy^1 \dots dy^n = \left| \det\left[\frac{\partial y^i}{\partial x^j}\right] \right| dx^1 \dots dx^n. \tag{9.29}$$

If g_{ij} and $\tilde{g}_{kl}$ represent the metric in the x and y coordinates, respectively, the chain rule gives the relation

$$\tilde{g}_{kl} = g_{ij} \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^l}. \tag{9.30}$$

We can thus define the density factor,

$$\sqrt{g} := \sqrt{\det[g_{ij}]},$$

so that (9.30) gives

$$\sqrt{\tilde{g}} = \left| \det\left[\frac{\partial x^i}{\partial y^k}\right] \right| \sqrt{g}.$$

Then, by (9.29), the volume density defined locally by

$$dV := \sqrt{g}\, dx^1 \dots dx^n \tag{9.31}$$

is invariant under changes of coordinates.

The global integral of a compactly supported function $f \in C(M)$,

$$f \mapsto \int_M f\, dV,$$

is defined by using a partition of unity to subdivide into integrals over local coordinate patches. By the standard measure theory construction, the integral

defined for continuous functions with compact support can be extended to a unique Borel measure on M, which is also denoted by dV.

The simplest way to define the Laplacian is through an analog of Green's identity: the operator Δ on $C^\infty(M)$ should satisfy

$$\int_M g(\nabla u, \nabla v)\, dV = -\int_M u \Delta v\, dV. \tag{9.32}$$

for all $u, v \in C_0^\infty(M)$. If u, v are supported within a local coordinate patch U, this becomes

$$\int_U g^{ij} \partial_i u\, \partial_j v \sqrt{g}\, d^n x = -\int_U u \Delta v \sqrt{g}\, d^n x.$$

Thus, by integration by parts, the local coordinate expression for the Laplacian is

$$\Delta := \frac{1}{\sqrt{g}} \partial_i \big(\sqrt{g} g^{ij} \partial_j\big). \tag{9.33}$$

This generalization of the Laplacian to Riemannian manifolds was introduced by Eugenio Beltrami and is also called the Laplace–Beltrami operator. In the geometric context, a minus is often included on the right-hand side of (9.33) as part of the definition.

Example 9.20. Consider the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$. In Example 9.5 we worked out the metric in spherical coordinates

$$ds^2 = d\varphi^2 + \sin^2\varphi\, d\theta^2.$$

Thus $\sqrt{g} = \sin\varphi$, and (9.33) becomes

$$\Delta := \frac{1}{\sin\varphi} \partial_\varphi \big(\sin\varphi\, \partial_\varphi\big) + \frac{1}{\sin^2\varphi} \partial_\theta^2.$$

Note that this spherical Laplacian appeared in (7.44) in the context of separation of variables in $\mathbb{R}^3$. ◊

9.3.1 Green's Identity

The local definition (9.33) of the Laplacian was inspired by Green's identity in form (9.32). Let us now adapt the identity to the case of a Riemannian manifold with boundary.

Theorem 9.21 (Green's Identity). *Let $\overline{\Omega}$ be a compact manifold with boundary. For $u, v \in C^\infty(\overline{\Omega})$,*

$$\int_\Omega \Big[g(\nabla u, \nabla v) + u\Delta v\Big] dV = \int_{\partial\Omega} u\, \partial_\nu v\, dS,$$

where ∂_ν is the outward unit normal tangent vector to $\partial\Omega$.

Proof By the assumption that $\overline{\Omega}$ is compact, we can choose a finite atlas of coordinate charts $\{U_j\}_{j=1}^m$. Let $\{\chi_j\}$ be a corresponding partition of unity as given by Lemma 9.4. By linearity, the identity can then be decomposed as

$$\sum_{i,j=1}^m \int_\Omega \Big[g\big(\nabla(\chi_i u), \nabla(\chi_j v)\big) + (\chi_i u)\Delta(\chi_j v)\Big] dV$$
$$= \sum_{i,j=1}^m \int_{\partial\Omega} (\chi_i u)\, \partial_\nu(\chi_j v)\, dS.$$

It therefore suffices to prove the result for functions which are compactly supported within a single coordinate neighborhood. Without loss of generality, we can take this to be a boundary neighborhood, so we will consider a metric g defined on an open set $W \subset \overline{\mathbb{R}^n_+}$, with h the induced metric on $\partial\overline{\mathbb{R}^n_+} = \{x^n = 0\}$.

Note that the determinant of h_{ij} is the (n, n) determinant minor of g_{ij}. Therefore, by the cofactor formula for the inverse matrix,

$$\det[h_{ij}] = g^{nn} \det[g_{ij}].$$

The determinant factors appearing in the measures dV and dS are thus related by

$$\sqrt{g} = \big(g^{nn}\big)^{-\frac{1}{2}} \sqrt{h} \tag{9.34}$$

For two functions $u, v \in C_0^\infty(W)$, consider the integral

$$\int_W g(\nabla u, \nabla v)\, dV = \int_W g^{ij} \partial_i u\, \partial_j v \sqrt{g}\, d^n x.$$

Applying integration by parts to the derivatives of u gives

$$\begin{aligned}\int_W g^{ij} \partial_i u\, \partial_j v\, \sqrt{g}\, d^n x = &-\int_W u\, \partial_i\big[g^{ij}\sqrt{g}\, \partial_j v\big]\, d^n x \\ &- \int_{\{x^n=0\}} u g^{nj} \partial_j v \sqrt{g}\, d^{n-1}x\end{aligned} \tag{9.35}$$

We claim that the outward unit normal to $\{x^n = 0\}$ is given by

$$\partial_\nu = -\big(g^{nn}\big)^{-\frac{1}{2}} g^{nj} \partial_j. \tag{9.36}$$

To check this, note that the tangent space to $\{x^n = 0\}$ is the span of ∂_i for $i = 1, \dots, n-1$. Therefore, to show that ∂_ν is normal to the boundary, it suffices to observe that

$$g^{nj} g_{ji} = 0.$$

for $i \neq n$, because g^{ij} is the inverse matrix to g_{ij}. We check that ∂_ν is a unit vector by noting that

$$g^{ni} g_{ij} g^{nj} = g^{nn}.$$

Finally, since $g^{nn} > 0$, the x^n component of ∂_ν is negative, as required for an outward normal to $\{x^n \geq 0\}$.

By (9.36) and (9.34) we have

$$\begin{aligned} -\int_{\{x^n=0\}} u g^{nj} \partial_j v \sqrt{g}\, d^{n-1}x &= \int_{\{x^n=0\}} u\, \partial_\nu v \big(g^{nn}\big)^{\frac{1}{2}} \sqrt{g}\, d^{n-1}x \\ &= \int_{\{x^n=0\}} u\, \partial_\nu v\, dS. \end{aligned}$$

Applying this to (9.35), along with the local formula (9.33) for the Laplacian, gives the desired formula,

$$\int_W \Big[g(\nabla u, \nabla v) + u \Delta v \Big] dV = \int_{\{x^n=0\}} u\, \partial_\nu v\, dS.$$

□

9.4 Spectrum of a Compact Manifold

The spectral theory of compact Riemannian manifolds, with or without boundary, has many parallels to the theory developed for bounded open sets in $\mathbb{R}^n$ in Chapter 6. In this section we will show that the Dirichlet Laplacian on a compact manifold with boundary has discrete spectrum with eigenvalues accumulating at infinity.

Example 9.22. Let Γ denote a lattice in $\mathbb{R}^n$, i.e., a group generated by a linearly independent set of translation vectors $\{v_n, \dots, v_n\} \subset \mathbb{R}^n$. The quotient $X_\Gamma :=$

$\mathbb{R}^n/\Gamma$ is a compact manifold whose functions can be identified with functions on $\mathbb{R}^n$ which are periodic under Γ. The torus $\mathbb{T}^n$ introduced in Example 2.32 is a special case.

The eigenfunctions of the Laplacian on X_Γ can easily be written down in terms of the dual lattice. Define

$$\Gamma^* := \left\{ w \in \mathbb{R}^n : \; w \cdot v \in 2\pi\mathbb{Z} \text{ for all } v \in \Gamma \right\}.$$

For $k \in \Gamma^*$, the function $x \mapsto e^{ik\cdot x}$ is then periodic with respect to Γ. As in the case of $\mathbb{T}^n$, we can use the Stone–Weierstrass theorem to argue that such functions are dense. This yields

$$\sigma(-\Delta_{X_\Gamma}) = \left\{ |k|^2 : \; k \in \Gamma^* \right\}.$$

$\Diamond$

Example 9.23. For the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ the Laplacian $-\Delta_{\mathbb{S}^2}$ appeared in Example 9.20. In spherical coordinates, the eigenfunctions are the spherical harmonics $Y_l^m(\varphi, \theta)$ that appeared in the proof of Theorem 7.17,

$$Y_l^m(\varphi, \theta) := c_l^m e^{im\theta} P_l^m(\cos\varphi)$$

for $l \in \mathbb{N}_0$ and $m \in \{-l, \dots, l\}$, where P_l^m is the Legendre function. A sample eigenfunction is shown in Figure 9.4. The spectrum is

$$\sigma(-\Delta_{\mathbb{S}^2}) = \left\{ l(l+1) : \; l \in \mathbb{N}_0, \text{multiplicity} = 2l + 1 \right\}.$$

$\Diamond$

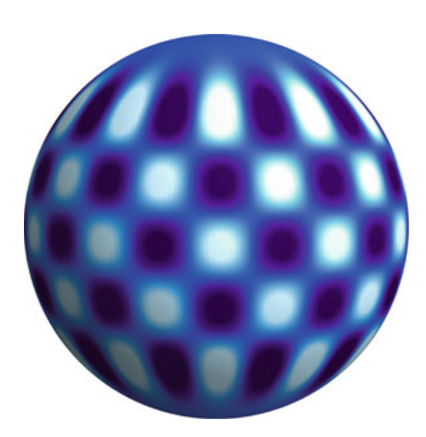


Fig. 9.4 A contour plot of the spherical harmonic Y_{11}^{15}

9.4.1 Dirichlet Eigenvalues

Let M be a Riemannian manifold, and $\Omega \subset M$ a relatively compact open set. As noted in Section 9.2, a general compact manifold with boundary can be realized in this way (with $\partial\Omega$ smooth). For the construction of the Dirichlet Laplacian we do not require any regularity of $\partial\Omega$.

As a differential operator, the Laplacian on $\overline{\Omega}$ is defined as in Section 9.3. We will consider the action of $-\Delta$ on $L^2(\Omega)$, the Hilbert space defined with respect to the volume measure dV induced by the metric. By Green's identity (Theorem 9.21), $-\Delta$ is a positive symmetric operator on $C_0^\infty(\Omega)$.

To define the Dirichlet extension of $-\Delta$ on $\overline{\Omega}$, we need to set up the H^1 Sobolev spaces. The space $H^1(\Omega)$ is defined as the set of functions $f \in L^2(\Omega)$ such that the restriction of f to each coordinate patch lies in $H^1(\mathbb{R}^n_+)$ when written in local coordinates. Lemma 6.13 shows that this definition is independent of the choice of coordinate patches.

For $u \in H^1(\Omega)$, the gradient ∇u is defined in the weak sense, by the local coordinate expressions $g^{ij}\partial_i u\, \partial_j$. We can thus define the inner product

$$\langle u, v\rangle_{H^1} := \langle u, v\rangle + \int_\Omega g(\nabla\overline{u}, \nabla v)\, dV. \tag{9.37}$$

The argument from Theorem 2.25 shows that $H^1(\Omega)$ is a Hilbert space with respect to this inner product.

As in Section 6.1.1, the subspace $H^1_0(\Omega)$ is defined as the closure of $C_0^\infty(\Omega)$ in $H^1(\Omega)$. Using Green's identity and approximation by functions in $C_0^\infty(\Omega)$ as in the derivation of (6.13), we can show that

$$\langle u, \psi\rangle_{H^1} = \langle u, -\Delta\psi\rangle + \langle u, \psi\rangle$$

for $u \in H^1_0(\Omega)$ and $\psi \in C_0^\infty(\Omega)$. Repeating the steps in the proof of Theorem 6.6 yields the construction of the Dirichlet Laplacian on $\overline{\Omega}$:

Theorem 9.24. *Let Ω be a relatively compact, open subset of a Riemannian manifold. The Laplacian $-\Delta$ is positive and self-adjoint on the domain*

$$\mathcal{D}(-\Delta) := \left\{u \in H^1_0(\Omega) :\ \Delta u \in L^2(\Omega)\right\},$$

where Δu is defined in the weak sense. This is the unique self-adjoint extension of $-\Delta$ on $C_0^\infty(\Omega)$ for which the domain is contained in $H^1_0(\Omega)$.

In the case of a compact manifold M without boundary, Theorem 9.24 implies that the Laplacian is essentially self-adjoint on $C^\infty(M)$. The arguments for compact resolvent when $\overline{\Omega}$ is compact can also be adapted from the Euclidean case.

Theorem 9.25. *For Ω a relatively compact, open subset of a Riemannian manifold, the Dirichlet Laplacian has compact resolvent and there exists an orthonormal basis for $L^2(\Omega)$ consisting of real-valued eigenfunctions $\psi_k \subset H_0^1(\Omega)$. The corresponding eigenvalues $\{\lambda_k\}$ are nonnegative and satisfy* $\lim \lambda_k = +\infty$.

Proof The argument is based on a straightforward generalization of Rellich's theorem (Theorem 6.9). Since $\overline{\Omega}$ is compact, there exists a finite coordinate atlas $\{U_j, \psi_j\}_{j=1}^m$. Let $\{\chi_j\}$ be a corresponding partition of unity. Suppose that $\{w_k\}$ is a bounded sequence in $H_0^1(\Omega)$. Then $\{\chi_j w_k\}$ is bounded in $H_0^1(U_j)$, and by applying Theorem 6.9 in each patch, we can pass to a subsequence such that $\chi_j w_k \to f_j \in L^2(U_j)$ for each j. It then follows that $w_k \to \sum_{j=1}^m f_j$ in $L^2(\Omega)$. This proves that, for a compact manifold with boundary, the inclusion

$$\iota : H_0^1(\Omega) \to L^2(\Omega) \tag{9.38}$$

is compact. As in the proof of Theorem 6.8, it follows that $(-\Delta+1)^{-1}$ is compact as an operator on L^2. The characterization of the spectrum follows from Theorem 4.21 (Hilbert–Schmidt). □

For a compact manifold without boundary, the lowest eigenvalue is $\lambda_0 = 0$, corresponding to the constant eigenfunction. If $\partial\Omega \neq \emptyset$, then the Dirichlet eigenvalues are strictly positive.

9.4.2 Regularity

For this subsection, we assume that $\overline{\Omega}$ is a compact manifold with boundary, so that $\partial\Omega$ is smooth. The Sobolev spaces $H^m(\Omega)$ are defined for $m \geq 2$ as the set of functions in $L^2(\Omega)$ for which the restriction to each coordinate patch lies in $H^m(\Omega)$, just as for $m = 1$. It is possible to define an inner product on $H^m(\Omega)$ that depends only on the metric, using covariant derivatives. However, a local-coordinate construction is sufficient for the purpose of regularity arguments.

Let $\{U_j, \psi_j)\}_{j=1}^q$ be a finite coordinate atlas for $\overline{\Omega}$, with $\{\chi_j\}_{j=1}^l$ a corresponding partition of unity as provided by Lemma 9.4. For $u, v \in H^m(\Omega)$, we define

$$\langle u, v\rangle_{H^m} := \sum_{j=1}^q \langle \chi_j u, \chi_j v\rangle_{H^m(U_j)}, \tag{9.39}$$

where each U_j is identified with the corresponding subset of $\mathbb{R}^n_+$, via the coordinate map, and $H^m(U_j)$ is defined as in Section 2.5. The corresponding norms are given by

$$\|u\|^2_{H^m} := \sum_{j=1}^q \|\chi_j u\|^2_{H^m(U_j)}.$$

Although the inner product and norm are dependent on the atlas and partition of unity, it is straightforward to check, using Lemma 6.13, that the resulting topology is independent of these choices.

We can deduce the local regularity of H^m functions on a manifold using Sobolev embedding (Theorem 2.26). Furthermore, Lemma 6.14 allows us to extend these results across a smooth boundary, yielding the following:

Theorem 9.26 (Sobolev Embedding). *Suppose $\overline{\Omega}$ is a compact Riemannian manifold with boundary. For $m > k + n/2$, $H^m(\Omega) \subset C^k(\overline{\Omega})$ for $m > k + n/2$.*

To establish the smoothness of eigenfunctions, we need a basic elliptic regularity result. For a solution of $-\Delta u = f$, this relates the Sobolev regularity of u to that of f.

Theorem 9.27 (Elliptic Regularity). *Let $\overline{\Omega}$ be a compact Riemannian manifold with boundary, with $-\Delta$ be the Dirichlet Laplacian defined in Section 9.3. If $u \in \mathcal{D}(\Delta)$ and $\Delta u \in H^m(\Omega)$ for some $m \in \mathbb{N}_0$, then $u \in H^{m+2}(\Omega)$, with*

$$\|u\|_{H^{m+2}} \le C\big(\|\Delta u\|_{H^m} + \|u\|\big),$$

where C depends only on Ω and m.

For operators on $\mathbb{R}^n$, Theorem 9.27 is a standard fact from PDE theory. Elliptic regularity is easily extended to manifolds, since the regularity issues are local. The details are somewhat technical, however, so we defer the proof to the Appendix A.4.

Corollary 9.28. *If ψ is an eigenfunction of the Dirichlet Laplacian on a compact manifold with boundary $\overline{\Omega}$, then $\psi \in C^\infty(\overline{\Omega})$.*

Proof Suppose that $\psi \in \mathcal{D}(L)$ with $-\Delta\psi = \lambda\psi$ for some $\lambda > 0$. This implies that $\lambda\psi \in H_0^1(\Omega)$, so Theorem 9.27 immediately gives $\psi \in H^3(\Omega)$. Then $\Delta\psi \in H^3(\Omega)$, which implies $\psi \in H^5(\Omega)$, etc. By induction, $\psi \in H^m(\Omega)$ for all m. Therefore, $\psi \in C^\infty(\overline{\Omega})$ by Theorem 9.26. □

In the case of a relatively compact subset $\Omega \subset M$, with no regularity assumptions on $\partial\Omega$, we can apply the same arguments in the interior using cutoffs. This gives the interior regularity $\psi \in C^\infty(\Omega)$ for an eigenfunction ψ in the general case.

9.5 Heat Equation

The Laplacian appears in the spatial component of the fundamental evolution equations, such as the heat and wave equations. We can thus naturally adapt these equations to manifolds simply by replacing the Euclidean version of the Laplacian with its Riemannian generalization (9.33). Analyzing such equations helps us to understand the interplay between geometry and physics. Furthermore, the physical evolution equations provide valuable tools for establishing the link between the

spectrum of the Laplacian and other geometric properties such as volume and curvature.

In this section, we will focus on the *heat equation*,

$$\partial_t u - \Delta u = 0, \tag{9.40}$$

on a compact Riemannian manifold M. As an application of our analysis of the heat kernel, we will derive the Weyl asymptotic formula for the eigenvalues.

Under the initial condition $u|_{t=0} = f$, (9.40) has the formal solution

$$u = e^{t\Delta} f. \tag{9.41}$$

Since $-\Delta$ is self-adjoint and positive, by Theorem 9.24, the *heat operator* $e^{t\Delta}$ is defined as a bounded operator on $L^2(M)$ for $t > 0$, by the functional calculus. The right-hand side of (9.41) is thus well defined, and we need to only check that it yields a solution.

Theorem 9.29. *Let M be a compact Riemannian manifold. For $f \in L^2(M)$, define $u(t, \cdot) := e^{t\Delta} f$ for $t > 0$. Then*

(a) *$u \in C^\infty(\mathbb{R}_+ \times M)$,*
(b) *u satisfies (9.40),*
(c) *$u(t, \cdot) \to f$ in the L^2 sense as $t \to 0^+$.*

Moreover, a solution with $u(t, \cdot) \in L^2(M)$ is uniquely determined by these conditions.

Proof The operator $\Delta^k e^{t\Delta}$ is bounded on $L^2(M)$ for $t > 0$, by the functional calculus, because the function $x \mapsto x^k e^{-tx}$ is bounded on $\mathbb{R}_+$. Hence,

$$\Delta^k u(t, \cdot) \in L^2(M), \tag{9.42}$$

for all $k \in \mathbb{N}$ and $t > 0$.

The next step is to check that $e^{t\Delta}$ is differentiable with respect to t in the operator topology. For $t < 0$,

$$\lim_{h\to 0} \frac{e^{-(t+h)x} - e^{-tx}}{h} = xe^{-tx},$$

and the mean value theorem shows that the convergence is uniform for $x \in [0, \infty)$. Therefore, by the functional calculus,

$$\lim_{h\to 0} \frac{1}{h}\Big[e^{(t+h)\Delta} - e^{t\Delta}\Big] = \Delta e^{t\Delta},$$

with convergence in operator norm. It follows that u is weakly differentiable with respect to t, and that the equation

$$\partial_t u = \Delta u \tag{9.43}$$

holds in the sense of weak derivatives. By (9.42), this shows that $\partial_t^k u(t, \cdot) \in L^2(M)$ for all $k \in \mathbb{N}$.

Smoothness of u can now be deduced from elliptic regularity. For $T > 0$, let $\chi \in C_0^\infty(0, T)$. The product manifold $\overline{\Omega} = [0, T] \times M$ is a compact manifold with boundary. For the metric $dt^2 + g$, the Laplacian is $\partial_t^2 + \Delta$, and by (9.42) and (9.43) we have

$$(\partial_t^2 + \Delta)^k (\chi u) \in L^2(\Omega)$$

for all $k \in \mathbb{N}$. Therefore, $\chi u \in C^\infty(\mathbb{R}_+ \times M)$ by Theorems 9.26 and 9.27, which proves (a).

The smoothness of u implies in particular that (9.43) holds in the classical sense, proving (b). From the functional calculus, the pointwise convergence $e^{tx} \to 1$ as $t \to 0^+$ implies that

$$\lim_{t \to 0^+} e^{t\Delta} = I$$

in the strong operator sense (see Example 2.12). Thus $u(t, \cdot) \to f$ in $L^2(M)$, which establishes (c).

Finally, to prove uniqueness, suppose u_1 and u_2 satisfy (a), (b), and (c), with $u_j(t, \cdot) \in L^2(M)$. Then $v := u_1 - u_2$ also solves the heat equation, implying that

$$\begin{aligned} \frac{d}{dt} \|v(t, \cdot)\|^2 &= \langle \partial_t v, v \rangle + \langle v, \partial_t v \rangle \\ &= \langle v, \Delta v \rangle \\ &\leq 0. \end{aligned}$$

Since $v(t, 0) \to 0$ in the L^2 sense as $t \to 0^+$, this shows that $v = 0$. □

Our main goal in this section will be to construct an integral kernel for the heat operator and then use its asymptotics as $t \to 0$ to derive the Weyl formula. The kernel construction is based on the Euclidean heat operator, so we first review that case.

Example 9.30. Let $-\Delta$ be the Laplacian on $\mathbb{R}^n$, with domain $H^2(\mathbb{R}^n)$. The heat operator $e^{t\Delta}$ is defined by the functional calculus and can be written explicitly using the Fourier transform. For $f \in C_0^\infty(\mathbb{R}^n)$,

$$\begin{aligned} e^{t\Delta} f(x) &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{ix\cdot\xi - t|\xi|^2} \hat{f}(\xi)\, d^n\xi \\ &= (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi - t|\xi|^2} f(y)\, d^n y\, d^n\xi. \end{aligned}$$

By Fubini's theorem, we can take the ξ integral first,

$$\int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi - t|\xi|^2}\, d^n\xi = \left(\frac{\pi}{t}\right)^{n/2} e^{-|x-y|^2/4t}.$$

Hence, with the integral kernel

$$\Psi(t; r) = (4\pi t)^{-n/2} e^{-r^2/4t}, \tag{9.44}$$

the heat operator can be written as

$$e^{t\Delta} f(x) = \int_{\mathbb{R}^n} \Psi(t; |x-y|) f(y)\, d^n y, \tag{9.45}$$

for $f \in C_0^\infty(\mathbb{R}^n)$.

Using the explicit integral kernel, we can see that the convergence as $t \to 0$ is actually better than expected from the functional calculus. In (9.45), the change of variables to $y = x + w\sqrt{t}$ yields

$$\int_{\mathbb{R}^n} \Psi(t, |x-y|) f(y)\, d^n y = (4\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-|w|^2/4} f(x + w\sqrt{t})\, d^n w.$$

For f continuous, it follows that

$$\lim_{t\to 0^+} \int_{\mathbb{R}^n} \Psi(t, |x-y|) f(y)\, d^n y = f(x), \tag{9.46}$$

uniformly on compact sets. ♢

9.5.1 Maximum Principle

A solution of the heat equation represents a distribution of temperatures as a function of time. Intuitively, since heat flows from warmer regions to colder regions, it should be impossible for a heat solution to take on a local maximum as a function of both time and position. The mathematical formulation of this idea will prove useful later on for error estimates in the heat kernel construction.

Theorem 9.31 (Maximum Principle for the Heat Equation). *Suppose M is a compact Riemannian manifold. If $u \in C^\infty(\mathbb{R}_+ \times M)$ satisfies the heat equation (9.40) and u extends continuously to $t = 0$, then*

$$\max_{t\ge 0, x\in M} u(t, x) = \max_{x\in M} u(0, x).$$

Proof For $\varepsilon > 0$, consider the shifted function $u^\varepsilon(t, x) := u(t, x) - \varepsilon t$. Assuming that u solves the heat equation, we have

$$\partial_t u^\varepsilon = \Delta u^\varepsilon - \varepsilon. \tag{9.47}$$

On the compact set $[0, T] \times M$, u^ε attains a maximum at some point (t_0, p_0). Suppose that $t_0 > 0$, which implies that

$$\partial_t u^\varepsilon(t_0, p_0) \geq 0$$

(with equality if $t_0 < T$). Then, by (9.47),

$$\Delta u^\varepsilon(t_0, p_0) \geq \varepsilon > 0. \tag{9.48}$$

Let $(x^1, \dots, x^n)$ be a set of geodesic polar coordinates centered at p_0. The fact that $u^\varepsilon(t_0, \cdot)$ has a local maximum at p_0 implies that

$$\partial_j^2 u^\varepsilon(t_0, p_0) \leq 0, \tag{9.49}$$

for each j. Hence $\Delta u^\varepsilon \leq 0$, contradicting (9.48).

We conclude that $t_0 = 0$, which yields

$$\max_{[0,T]\times M} u^\varepsilon = \max_{x\in M} u^\varepsilon(0, x).$$

Hence, by the definition of u^ε,

$$\max_{[0,T]\times M} u \leq \max_{x\in M} u(0, x) + \varepsilon T.$$

The claim is proven by taking first $\varepsilon \to 0$, and then $T \to \infty$. □

Note that the maximum principle applies also to $-u$, yielding an equivalent minimum principle,

$$\min_{t\geq 0, x\in M} u(t, x) = \min_{x\in M} u(0, x).$$

9.5.2 Heat Kernel

The relationship between heat propagation and geometry was first worked out in 1949, in the following characterization of the *heat kernel* of a Riemannian manifold [62].

Theorem 9.32 (Minakshisundaram–Pleijel). *Let M be a compact Riemannian manifold. There exists a function $H \in C^\infty(\mathbb{R}_+ \times M \times M)$ such that*

$$e^{t\Delta} f(x) = \int_M H(t; x, y) f(y)\, dV(y), \tag{9.50}$$

for $f \in L^2(M)$. If f is continuous, then

$$\lim_{t\to 0^+} \int_M H(t; x, y) f(y)\, dV(y) = f(x). \tag{9.51}$$

For each $x \in M$, there is a uniform asymptotic expansion as $t \to 0^+$,

$$H(t, x, x) \sim (4\pi t)^{-n/2} \sum_{j=0}^{\infty} \alpha_j(x) t^j, \tag{9.52}$$

where $\alpha_0 = 1$ and α_j depends only on the metric and its derivatives at x.

Theorem 9.32 can be extended to complete manifolds, with or without boundary. We restrict our attention to the compact case here to simplify the discussion. The proof is based on the construction of a *parametrix*, i.e., an approximate solution that captures the essential features of the true heat kernel. The model for the heat parametrix is based on the Euclidean heat kernel from Example 9.30.

Given a point $y \in M$, let g_{ij} denote the metric in a set of geodesic normal coordinates $x^1, \dots, x^n$ centered at y. In terms of the corresponding polar coordinates $(r, \omega) \in [0, r_0) \times \mathbb{S}^{n-1}$, we will write the volume density factor as

$$\varphi_y(r, \omega) := \sqrt{g}.$$

By Lemma 9.11, φ_y is smooth and satisfies

$$\varphi_y = 1 + O(r^2)$$

as $r \to 0$.

According to Lemma 9.12, the metric has the form $ds^2 = dr^2 + h$ in geodesic polar coordinates, where h denotes a family of metrics on $\mathbb{S}^{n-1}$ parametrized by r. Since the Jacobian of the change of coordinates from x^j to (r, ω) equals r^{n-1}, the volume density factor $\sqrt{h}$ is related to φ_y by

$$\sqrt{h} = r^{n-1}\varphi_y.$$

The Laplacian of a radial function $\psi(r)$ is thus given by

$$\begin{aligned}\Delta\psi &= \frac{1}{\sqrt{h}}\partial_r\Big(\sqrt{h}\,\partial_r\psi\Big)\\ &= \partial_r^2\psi + (n-1)r^{-1}\partial_r\psi + \partial_r(\log\varphi_y)\psi.\end{aligned} \tag{9.53}$$

Our first goal is to derive a local coordinate version of the heat expansion in polar coordinates. We start from the Euclidean heat kernel $\Psi(t;r)$ defined by (9.44). By identifying r with the geodesic polar coordinate centered at y, we can interpret Ψ as a function on $\mathbb{R}_+ \times B(y;r_0)$, where $r_0 := \operatorname{inj}(M)$. From (9.53) we compute that

$$(\partial_t - \Delta_x)\Psi(t;r) = \frac{r}{2t}\partial_r(\log\varphi_y)\,\Psi(t;r), \tag{9.54}$$

where $r := \operatorname{dist}(x,y)$.

Equation (9.54) serves as the basis for the iterative construction of a parametrix. The argument given below could be applied to the case of a complete manifold, provided there is a lower bound on the injectivity radius. In that case, however, the error estimate would only be uniform on compact sets.

Lemma 9.33 (Heat Parametrix). *On a compact Riemannian manifold M, for each $k \in \mathbb{N}$, there exists a function $F_k \in C^\infty(\mathbb{R}_+ \times M \times M)$ with the following properties:*

(a) *F_k is an approximate solution of the heat equation, in the sense that*

$$(\partial_t - \Delta_x)F_k(t,x,y) = O(t^{k-n/2})$$

as $t \to 0$, uniformly for $(x,y) \in M \times M$.

(b) *For $f \in C^\infty(M)$,*

$$\lim_{t\to 0+}\int_M F_k(t,x,y)f(y)\,dV(y) = f(x),$$

uniformly for $x \in M$.

Proof We start by using the Euclidean heat kernel Ψ to make the local ansatz,

$$F_k(t,x,y) = \sum_{j=0}^k t^j u_j(x,y)\Psi(t;r),$$

for $r := \operatorname{dist}(x,y) < r_0$, where r_0 is the injectivity radius of M. The coefficients u_j are yet to be determined. From (9.54) we compute that

$$
\begin{aligned}
&(\partial_t - \Delta_x) \sum_{j=0}^{k} t^j u_j \Psi(t; r) \\
&= \sum_{j=0}^{k} \Big[j u_j t^{j-1} + \frac{r}{2t} \partial_r (\log \varphi_y) u_j t^j - t^j \Delta_x u_j - \frac{r}{t} (\partial_r u_j) t^j \Big] \Psi \\
&= \Big[\frac{r}{2} \partial_r (\log \varphi_y) u_0 - r \partial_r u_0 \Big] t^{-1} \Psi \\
&\quad + \sum_{j=1}^{k} \Big[j u_j + \frac{r}{2} \partial_r (\log \varphi_y) u_j - \Delta_x u_{j-1} - r \partial_r u_j \Big] t^{j-1} \Psi - (\Delta_x u_k) t^k \Psi.
\end{aligned}
$$

To cancel the leading t^{-1} term, we set

$$
u_0 = \varphi_y^{-\frac{1}{2}}. \tag{9.55}
$$

Setting the coefficient of the term of order t^{j-1} equal to zero then gives a recursive relation,

$$
r \partial_r \Big[r^j \varphi_y^{\frac{1}{2}} u_j \Big] = r^j \varphi_y^{\frac{1}{2}} \Delta_x u_{j-1}.
$$

This can be solved by an integral formula,

$$
u_j(r\omega, y) = r^{-j} \varphi_y(r, \omega)^{-\frac{1}{2}} \int_0^r s^{j-1} \varphi_y(s, \omega)^{-\frac{1}{2}} \Delta_x u_{j-1}(s\omega, y)\, ds, \tag{9.56}
$$

for $j = 1, \ldots, k$. The smooth dependence of ODE solutions on initial data from Theorem 9.8 implies that $\varphi_y(x)$ is smooth in both variables. It thus follows from (9.56) that the coefficients u_j are smooth functions.

The local coefficient formulas (9.55) and (9.56) are only valid for $\mathrm{dist}(x, y) < r_0$. To define the parametrix globally, we introduce a cutoff $\chi \in C^\infty[0, \infty)$, with $\chi(r) = 1$ for $r \leq r_0/2$ and $\chi(r) = 0$ for $r \geq 3r_0/4$. For each $x, y \in M$ we define

$$
F_k(t, x, y) := \chi(\mathrm{dist}(x, y)) \sum_{j=0}^{k} t^j u_j(x, y) \Psi(t, \mathrm{dist}(x, y)). \tag{9.57}
$$

By the construction of u_j,

$$
(\partial_t - \Delta) F_k = -(\chi \circ d) t^k (\Delta u_k) \Psi - [\Delta, \chi \circ d] \sum_{j=0}^{k} t^j u_j \Psi. \tag{9.58}
$$

Since Δu_k is smooth and depends only on the metric and its derivatives, and $t^{n/2}\Psi$ is continuous, the first term is uniformly $O(t^{k-n/2})$. For the second term, the operator $[\Delta, \chi \circ d]$ restricts the support of $\Psi(t,r)$ to $r \in [r_0/2, 3r_0/4]$. For r in this range, Ψ and its derivatives satisfy a uniform $O(t^N)$ estimate for any $N > 0$. This proves (a).

To prove (b), a simple scaling argument shows that only the $j = 0$ term contributes in the limit,

$$\lim_{t\to 0+} \int_M F_k(t,x,y) f(y)\, dV(y)$$
$$= \lim_{t\to 0+} \int_M \chi(\mathrm{dist}(x,y)) u_0(x,y) \Psi(t, \mathrm{dist}(x,y)) f(y)\, dV(y).$$

If we write the integral in geodesic polar coordinates centered at x, then the measure becomes $dV(y) = \varphi(y,x)\, d^n y$. If f is continuous, then by (9.46) and the fact that $u_0(x,x) = 1$, we have

$$\lim_{t\to 0+} \int_M F_k(t,x,y) f(y)\, dV(y) = f(x).$$

The limit is uniform in x because f is uniformly continuous by the compactness of M. □

We are now prepared to prove Theorem 9.32, using the parametrix provided by Lemma 9.33 as a starting point.

Proof of Theorem 9.32 Let F_k be the heat parametrix constructed as in Lemma 9.33, for some $k > n/2$. Let us write the error term from (9.58) as $R_k(t,x,y)$, so that

$$(\partial_t - \Delta) F_k = R_k. \tag{9.59}$$

Our goal is to modify F_k to remove this error term, without affecting the behavior of F_k as $t \to 0$. Since we have at our disposal the heat operator $e^{t\Delta}$ provided by the spectral theorem, the most direct approach is to solve the inhomogeneous heat equation

$$(\partial_t - \Delta) Q_k = R_k,$$

subject to the initial condition $Q_k|_{t=0} = 0$.

This inhomogeneous problem can be solved using *Duhamel's principle*, which gives

$$Q_k(t,x,y) := \int_0^t e^{(t-s)\Delta} R_k(s,x,y)\, ds. \tag{9.60}$$

Note that R_k is smooth for $t > 0$ and $R_k = O(t^{k-n/2})$ uniformly as $t \to 0^+$. Using these facts, we can justify differentiation under the integral to obtain

$$\begin{aligned}\partial_t Q_k(t,x,y) &= R_k(t,x,y) + \int_0^t \Delta e^{(t-s)\Delta} R_k(s,x,y)\,ds \\ &= R_k(t,x,y) + \Delta Q_k(t,x,y).\end{aligned} \tag{9.61}$$

We now set $H := F_k - Q_k$, and observe that $H \in C^\infty(\mathbb{R}_+ \times M \times M)$ and

$$(\partial_t - \Delta)H = 0 \tag{9.62}$$

by (9.59) and (9.61). Given $f \in C^\infty(M)$, set

$$u(t,x) := \int_M H(t,x,y) f(y)\,dV(y) \tag{9.63}$$

for $t > 0$. Since H is smooth in the variables (t,x) by construction, differentiation under the integral in (9.63) is easy to justify. It follows u is smooth and satisfies $(\partial_t - \Delta)u = 0$ by (9.62).

To study the limit $t \to 0$ we need to estimate Q_k in this limit. The maximum principle (Theorem 9.31) implies that for fixed s,

$$\max_{t \geq s, x, y \in M} \left| e^{(t-s)\Delta} R_k(s,x,y) \right| = \max_{x,y\in M} \left| R_k(s,x,y) \right|.$$

Since $R_k(s,x,y)$ is uniformly $O(s^{k-n/2})$, by Lemma 9.33, we can estimate the integral in (9.60) to obtain

$$Q_k(t,\cdot,\cdot) = O(t^{k-n/2+1}), \tag{9.64}$$

uniformly as $t \to 0$.

Using (9.64), along with part (b) of Lemma 9.33, we conclude that

$$u(t,\cdot) \to f,$$

uniformly as $t \to 0^+$. Therefore u satisfies the conditions of Theorem 9.29. By uniqueness, this means that

$$e^{t\Delta} f(x) = \int_M H(t,x,y) f(y)\,dV(y), \tag{9.65}$$

for $f \in C^\infty(M)$. Since both sides represent the action of bounded operators, the equality extends to $f \in L^2(M)$.

The asymptotic expansion of $H(t, x, x)$ follows from the construction of F_k in Lemma 9.33. By (9.57), for each k we have

$$H(t, x, x) = (4\pi t)^{-n/2} \sum_{j=0}^{k} t^j u_j(x, x) + Q_k(t, x, x). \tag{9.66}$$

The uniform estimate on the remainder term follows from (9.64). □

9.5.3 *Spectral Applications*

Let M be a compact Riemannian manifold. By Theorem 9.25, there is an orthonormal basis $\{\psi_j\}$ for $L^2(M)$, consisting of eigenvalues of $-\Delta$ with corresponding eigenvalues

$$0 = \lambda_0 < \lambda_1 < \cdots \to \infty.$$

The existence of the smooth heat kernel shows that the heat operator $e^{t\Delta}$ is Hilbert–Schmidt.

By adapting the proof of Mercer's theorem (Theorem 4.23), we can deduce that

$$H(t, x, y) = \sum_{k=0}^{\infty} e^{-\lambda_k t} \psi_k(x) \overline{\psi_k(y)},$$

with uniform convergence on $[t_0, \infty) \times M \times M$ for $t_0 > 0$. It then follows from the proof of Theorem 4.24 that $e^{t\Delta}$ is trace-class for $t > 0$, with

$$\operatorname{tr}\bigl(e^{t\Delta}\bigr) = \int_M H(t, x, x)\, dV(x).$$

Since the asymptotic expansion in Theorem 9.32 is uniform, it can be integrated over M to obtain

$$\int_M H(t, x, x)\, dV(x) \sim (4\pi t)^{-n/2} \sum_{j=0}^{\infty} a_j t^j,$$

where

$$a_j := \int_M \alpha_j \, dV.$$

The coefficients $\{a_j\}$ are called the *heat invariants* of M. They can be computed explicitly, in terms of Riemannian curvature, but we will not get into that here, except to note that since $\alpha_0 = 1$,

$$a_0 = \operatorname{vol}(M).$$

Writing the trace as a sum over the eigenvalues yields the following:

Theorem 9.34. *The heat invariants of a compact Riemannian manifold are determined by the spectrum, through the asymptotic expansion*

$$\sum_{k=0}^{\infty} e^{-\lambda_k t} \sim (4\pi t)^{-n/2} \sum_{j=0}^{\infty} a_j t^j,$$

as $t \to 0^+$.

The leading term in the heat expansion gives the Weyl asymptotic for eigenfunctions on a compact manifold. Define the counting function

$$N_M(\tau) := \#\{\lambda_k \le \tau\}.$$

Corollary 9.35 (Weyl's Asymptotic Formula). *The eigenvalues of a compact Riemannian manifold satisfy*

$$N_M(\tau) \sim (2\pi)^{-n} \omega_n \operatorname{vol}(M) \tau^{\frac{n}{2}} \tag{9.67}$$

as $t \to \infty$, *where* ω_n *is the volume of the unit ball in* $\mathbb{R}^n$, (6.50). *Equivalently, if the eigenvalues are arranged in increasing order,*

$$\lambda_k \sim (2\pi)^2 \left(\frac{k}{\omega_n \operatorname{vol}(M)} \right)^{2/n}$$

as $k \to \infty$.

Proof The heat trace can be written as a Stieljes integral

$$\sum_{k=0}^{\infty} e^{-\lambda_k t} = \int_0^{\infty} e^{-tx} dN_M(x).$$

By Karamata's result (Theorem 6.33), the asymptotic from Theorem 9.34 implies that

$$N_M(\tau) \sim (4\pi)^{-\frac{n}{2}} \frac{\operatorname{vol}(M)}{\Gamma(n/2+1)} \tau^{\frac{n}{2}}.$$

□

Example 9.36. Let us consider the two explicit cases from Examples 9.22 and 9.23. For the quotient $X_\Gamma := \mathbb{R}^n/\Gamma$, the counting function is given by

$$N_{X_\Gamma}(\tau) = \#\left\{ k \in \Gamma^* : |k| \le \sqrt{\tau} \right\}.$$

The number of lattice points in a ball of radius $\sqrt{\tau}$ can be estimated as in Lemma 6.21, to deduce that

$$N_{X_\Gamma}(\tau) \sim \frac{\mathrm{vol}(B(0;\sqrt{\tau}))}{\mathrm{vol}(\mathbb{R}^n/\Gamma^*)}. \tag{9.68}$$

If A is the linear transformation mapping the standard basis for $\mathbb{R}^n$ to a set of generators $\{v_1, \ldots, v_n\}$ for Γ, then

$$\mathrm{vol}(X_\Gamma) = \det A.$$

By the definition of the dual lattice, $2\pi(A^t)^{-1}$ maps the standard basis to a set of generators for Γ^*. Thus we have

$$\begin{aligned}\mathrm{vol}(\mathbb{R}^n/\Gamma^*) &= \det(2\pi A^{-1})\\ &= \frac{(2\pi)^n}{\mathrm{vol}(X_\Gamma)}.\end{aligned}$$

From (9.68) we thus obtain

$$N_{X_\Gamma}(\tau) \sim (2\pi)^{-n}\,\mathrm{vol}(X_\Gamma)\omega_n \tau^{\frac{n}{2}},$$

in agreement with (9.67).

For the unit sphere $\mathbb{S}^2$, the eigenvalues are given by $l(l+1)$ for $l \in \mathbb{N}_0$, each with multiplicity $2l+1$. The number of eigenvalues with $0 \le l \le q$ is given by

$$\sum_{l=0}^{q}(2l+1) = (q+1)^2.$$

Since $(q+1)^2 \sim q(q+1)$, this gives

$$N_{\mathbb{S}^2}(\tau) \sim \tau.$$

Since $\mathrm{vol}(\mathbb{S}^2) = 4\pi$ and $\omega_2 = \pi$, this matches (9.67). ♢

9.6 Wave Propagation on Compact Manifolds

We turn next to the analysis of the *wave equation*,

$$\partial_t^2 u - \Delta u = 0, \tag{9.69}$$

on a Riemannian manifold. In this case we will focus on compact manifolds with boundary.

As with the heat equation, the functional calculus can be used to construct weak solutions of the wave equation. The functions $\cos(tx)$ and $\sin(tx)/x$ are smooth and bounded for $t, x \in \mathbb{R}$. We can thus define two families of *wave operators* on $L^2(\Omega)$, by applying these functions to $\sqrt{-\Delta}$. Formally at least, the function

$$u(t,\cdot) := \cos(t\sqrt{-\Delta})u_0 + \frac{\sin(t\sqrt{-\Delta})}{\sqrt{-\Delta}} v_0 \tag{9.70}$$

solves (9.69) for the initial conditions

$$u(0,\cdot) = u_0, \qquad \partial_t u(0,\cdot) = v_0. \tag{9.71}$$

We will show that u is a weak solution of (9.69), in the sense that $u \in C^2(\mathbb{R}, L^2(\Omega))$, $u(t,\cdot) \in \mathcal{D}(\Delta)$ for all t, and u satisfies $\partial_t^2 u = \Delta u$. The space $C^m(\mathbb{R}, L^2(\Omega))$ is defined as the set of continuous functions $\mathbb{R} \to L^2(\Omega)$ which admit m continuous derivatives defined as L^2-limits of difference quotients. The condition that $u \in C^2(\mathbb{R}, L^2(\Omega))$ implies in particular that $\partial_t^2 u$ exists as a weak derivative.

Theorem 9.37. *Let $\overline{\Omega}$ be a compact manifold with boundary, and $-\Delta$ the Dirichlet Laplacian as defined in Section 9.3. Given $u_0 \in \mathcal{D}(\Delta)$, $v_0 \in H_0^1(\Omega)$, the function u defined by* (9.70) *is the unique weak solution of the wave equation that satisfies the initial conditions* (9.71).

If u_0 and v_0 are smooth functions that vanish to all orders at $\partial\Omega$, then $u \in C^\infty(\mathbb{R} \times \overline{\Omega})$.

Proof Let $\{\psi_k\} \subset H_0^1(\Omega)$ be the eigenfunction basis for $L^2(\Omega)$ given in Theorem 9.25, with eigenvalues $\{\lambda_k\}$. Note that

$$\|\psi_k\|_{H^1}^2 = (1 + \lambda_k).$$

For $f \in L^2(\Omega)$, this implies that

$$\|f\|_{H^1}^2 = \sum_{k=1}^\infty (1 + \lambda_k) \left|\langle \psi_k, f\rangle\right|^2. \tag{9.72}$$

Since the basis consists of eigenfunctions, we also have

$$\|\Delta f\|^2 = \sum_{k=1}^\infty \lambda_k^2 \left|\langle \psi_k, f\rangle\right|^2.$$

In conjunction with (9.72) this shows that, for $f \in L^2(\Omega)$,

$$f \in \mathcal{D}(\Delta) \quad \Longleftrightarrow \quad \sum_{k=1}^\infty \lambda_k^2 \left|\langle \psi_k, f\rangle\right|^2 < \infty. \tag{9.73}$$

Given $u_0 \in \mathcal{D}(-\Delta)$ and $v_0 \in H_0^1(\Omega)$, define the coefficients

$$a_k := \langle u_0, \psi_k \rangle, \qquad b_k := \langle v_0, \psi_k \rangle.$$

The function u defined by (9.70) is then given by

$$u = \sum_{k=1}^{\infty} \left[\cos(t\sqrt{\lambda_k}) a_k \psi_k + \frac{\sin(t\sqrt{\lambda_k})}{\sqrt{\lambda_k}} b_k \psi_k \right]. \tag{9.74}$$

By (9.72) and (9.73), the coefficients satisfy

$$\sum_{k=1}^{\infty} \lambda_k^2 |a_k|^2 < \infty, \qquad \sum_{k=1}^{\infty} \lambda_k |b_k|^2 < \infty. \tag{9.75}$$

It thus follows from (9.73) that $u(t, \cdot) \in \mathcal{D}(\Delta)$.

The estimates (9.75) also justify taking L^2 derivatives of u with respect to t by differentiating (9.74) term by term. For example, the derivative of the first term is

$$\partial_t \sum_{k=1}^{\infty} \cos(t\sqrt{\lambda_k}) a_k \psi_k = \lim_{h \to 0} \sum_{k=1}^{\infty} \left[\frac{\cos((t+h)\sqrt{\lambda_k}) - \cos(t\sqrt{\lambda_k})}{h} \right] a_k \psi_k.$$

By the mean value theorem, the expression in brackets is bounded by $\sqrt{\lambda_k}$, uniformly in t and h. Therefore, by (9.75), we can apply the dominated convergence theorem to take the limit $h \to 0$ inside the sum, in the L^2 sense. This yields

$$\partial_t \sum_{k=1}^{\infty} \cos(t\sqrt{\lambda_k}) a_k \psi_k = -\sum_{k=1}^{\infty} \sqrt{\lambda_k} \sin(t\sqrt{\lambda_k}) a_k \psi_k.$$

By the same argument, we can see that $\partial_t^2 u$ exists as a weak L^2 derivative, given by

$$\partial_t^2 u = -\sum_{k=1}^{\infty} \left[\lambda_k \cos(t\sqrt{\lambda_k}) a_k \psi_k + \sqrt{\lambda_k} \sin(t\sqrt{\lambda_k}) b_k \psi_k \right].$$

This shows that $u \in C^2(\mathbb{R}, L^2(\Omega))$ and u is a weak solution of (9.69).

To prove that the weak solution is uniquely determined by initial conditions, we use an argument based on conservation of energy. The total energy of the wave solution at time t is defined as

$$E(t) := \frac{1}{2} \left\| \partial_t u(t, \cdot) \right\|_{L^2(\Omega)}^2 + \frac{1}{2} \left\| \nabla u(t, \cdot) \right\|_{L^2(\Omega)}^2. \tag{9.76}$$

If u is a weak solution of (9.69), then E is finite and satisfies

$$\begin{aligned}\partial_t E &= \mathrm{Re}\langle \partial_t u, \partial_t^2 u\rangle + \mathrm{Re}\langle \nabla u, \partial_t \nabla u\rangle \\ &= \mathrm{Re}\langle \partial_t u, \Delta u\rangle - \mathrm{Re}\langle u, \partial_t \Delta u\rangle \\ &= 0.\end{aligned}$$

The total energy is thus independent of t. If u is a solution with vanishing initial data u_0 and v_0, then $E = 0$ for all $t \in \mathbb{R}$, implying that u is constant and therefore identically 0. By linearity, it follows that the solution is uniquely determined by u_0 and v_0.

For the additional regularity claim, suppose that u_0, v_0 are smooth and vanish to infinite order at $\partial\Omega$. Then $\Delta^l u_0$ and $\Delta^l u_0$ are contained in $\mathcal{D}(\Delta)$ for all $l \in \mathbb{N}_0$. It then follows from (9.73) that

$$\sum_{k=1}^{\infty} \lambda_k^{2l} |a_k|^2 < \infty, \qquad \sum_{k=1}^{\infty} \lambda_k^{2l} |b_k|^2 < \infty,$$

for all l. This implies, again by (9.73), that $\Delta^l u(t, \cdot) \in \mathcal{D}(\Delta)$ for all l. Hence $u(t, \cdot) \in H^m(\Omega)$ for all m by Theorem 9.27, and therefore $u(t, \cdot) \in C^\infty(\overline{\Omega})$ by Theorem 9.26. The same argument shows that the weak derivatives $\partial_t^m u$ exist for all m and are represented by smooth functions on $\overline{\Omega}$. It then follows from Lemma 2.22 that u is smooth as a function of t also. □

9.6.1 Propagation Speed

In the classical theory of wave propagation, Huygens' principle says that wave fronts travel at a fixed propagation speed, which is equal to 1 for the wave equation written in the form (9.69). Mathematically, this property is usually formulated as a statement about the rate at which the support of a solution will expand, which is called the *propagation speed*.

Because we have only demonstrated the existence of solutions of the wave equation on a compact manifold with boundary, we will restrict our attention to that context for this result. Generalizing this principle to non-compact manifolds is straightforward.

Theorem 9.38 (Finite Propagation Speed). *Let $\overline{\Omega}$ be a compact Riemannian manifold with boundary, and suppose u is a smooth solution of the wave equation as given by Theorem 9.37, with Dirichlet boundary conditions on $\partial\Omega$ and initial conditions $u_0, v_0 \in C_0^\infty(\Omega)$. If $K := \operatorname{supp} u_0 \cup \operatorname{supp} v_0$, then*

$$\operatorname{supp} u(\cdot, t) \subset \left\{x \in \overline{\Omega} : \operatorname{dist}(x, K) \le |t|\right\}. \tag{9.77}$$

Proof It suffices to consider the case when u is real-valued, by linearity. Also, since $u(x, -t)$ solves the wave equation with initial conditions u_0 and $-v_0$, we need to only consider $t \geq 0$. As noted in the remarks following Definition 9.2, we may assume that the manifold with boundary $\overline{\Omega}$ is a closed domain within a compact manifold M.

Given $p_0 \in \overline{\Omega}$, let $U_p(r_0)$ be a geodesic polar neighborhood in M, with radial coordinate r, as described in Theorem 9.15. The set $U_p(r_0)$ may extend outside the boundary of Ω into M. For $0 \leq t \leq r_0$ we set

$$\begin{aligned} W_p(t) &:= \left\{q \in \overline{\Omega} : \operatorname{dist}(p, q) \leq r_0 - t\right\} \\ &= \{r \leq r_0 - t\} \cap \Omega. \end{aligned}$$

We adapt the definition (9.76) of total energy, to write the energy of u contained with $W_p(t)$ as

$$\eta(t) := \int_{W_p(t)} \Big[(\partial_t u)^2 + \|\nabla u\|_g^2\Big] dV.$$

where $\|\nabla u\|_g^2 := g(\nabla u, \nabla u)$ and the integrand is evaluated at time t.

Because ∂_r is orthogonal to spheres of constant r, by the Lemma 9.12 (Gauss), the volume density dV decomposes as $dr \times dS$, where dS is the measure on $\partial W_p(t)$ induced by the restriction of g. Separating the radial and spherical integrals allows us to write

$$\eta(t) = \int_0^{r_0 - t} h(r, t)\, dr,$$

where

$$h(r, t) := \int_{\partial B(p;r) \cap \Omega} \Big[(\partial_t u)^2 + \|\nabla u\|_g^2\Big] dS.$$

The derivative is then

$$\eta'(t) = -h(r_0 - t, t) + \int_0^{r_0 - t} \partial_t h(r, t)\, dr. \tag{9.78}$$

Using the fact that u satisfies the wave equation (9.69), we have

$$\begin{aligned} \partial_t h(r, t) &= 2 \int_{\partial B(p;r) \cap \Omega} \Big[(\partial_t^2 u)(\partial_t u) + g(\partial_t \nabla u, \nabla u)\Big] dS \\ &= 2 \int_{\partial B(p;r) \cap \Omega} \Big[(\partial_t u)(\Delta u) + g(\partial_t \nabla u, \nabla u)\Big] dS. \end{aligned}$$

Thus

$$\int_0^{r_0-t} \partial_t h(r,t)\, dr = 2\int_{W_p(t)} \Big[(\partial_t u)(\Delta u) + g(\partial_t \nabla u, \nabla u)\Big] dV. \tag{9.79}$$

Green's identity (Theorem 9.21) is easily extended to the case of piecewise smooth boundary, yielding

$$\int_{W_p(t)} \Big[(\partial_t u)\Delta u + g(\partial_t \nabla u, \nabla u)\Big] dV = \int_{\partial W_p(t)} (\partial_t u)\partial_\nu u\, dS.$$

The boundary of $W_p(t)$ has two components, $\partial B(p; r_0 - t) \cap \Omega$ and $B(p; r_0 - t) \cap \partial \Omega$. However, since $u|_{\partial\Omega} = 0$, the integrand vanishes on the latter. Therefore (9.79) can be reduced to

$$\int_0^{r_0-t} \partial_t h(r,t)\, dr = 2\int_{\partial B(p;r_0-t)\cap\Omega} (\partial_t u)\partial_r u\, dS. \tag{9.80}$$

Since ∂_r is a unit vector, $|\partial_r u| \leq \|\nabla u\|_g$, and so plugging (9.80) back into (9.78) and using the definition of $h(r_0 - t, t)$ give

$$\eta'(t) \leq -\int_{\partial B(p;r_0-t)\cap\Omega} \Big[(\partial_t u)^2 - 2(\partial_t u)\|\nabla u\|_g + \|\nabla u\|_g^2\Big] dS.$$

The integrand is nonnegative, implying that $\eta'(t) \leq 0$.

Suppose now that u and $\partial_t u$ vanish on $W_p(0)$ at time $t = 0$, implying that $\eta(0) = 0$. Since $\eta'(t) \leq 0$, this means that $\eta(t) = 0$ for all $t \in [0, r_0]$. Therefore u vanishes on the set $\{(x,t) : x \in W_p(t)\}$. In particular, $u(p, r_0) = 0$.

Since $\overline{\Omega}$ is compact, by Lemma 9.18 there exists $r_0 > 0$ such that each $U_p(r_0)$ is a geodesic polar neighborhood in M for each $p \in \overline{\Omega}$. The argument given above then applies to all points p with $\mathrm{dist}(p, K) \geq r_0$, which proves that (9.77) holds for $0 \leq t \leq r_0$. The full claim then follows by iterating this argument from later starting times. □

Note that Theorem 9.38 essentially says that the time taken for a signal from outside of K to reach an interior point is bounded by the distance from the interior point to ∂K, i.e., the propagation speed is at most one.

9.7 Complete Manifolds and Essential Self-adjointness

In this section, we consider the problem of essential self-adjointness of the Laplacian on $C_0^\infty(M)$, where M is a complete, non-compact Riemannian manifold. The one such case we have dealt with so far is Euclidean $\mathbb{R}^n$, for which essential

self-adjointness was demonstrated in Example 3.26. That argument was based on the Fourier transform and does not generalize to manifolds.

The example of a bounded open subset $\Omega \subset \mathbb{R}^n$, considered in Section 9.4, shows that essential self-adjointness may fail in the non-compact setting, in cases where boundary conditions are needed. The crucial property which distinguishes $\mathbb{R}^n$ from a bounded open subset is completeness.

Theorem 9.39. *For a complete Riemannian manifold M, the Laplacian is essentially self-adjoint on $C_0^\infty(M)$.*

This result, proven independently by Gaffney [33] and Roelcke [75], is fundamental for the spectral analysis of non-compact manifolds. It implies the self-adjoint extension of $-\Delta$ is uniquely defined, and therefore its spectral theory yields geometric invariants. The existence of a self-adjoint extension also gives access to the functional calculus for $-\Delta$, an essential tool in the analysis of PDE on M.

On a more practical level, Theorem 9.39 allows us to analyze the spectrum using smooth, compactly supported test functions. For example, it shows that the location of the bottom of the spectrum is given by

$$\inf \sigma(-\Delta) = \inf_{f \in C_0^\infty(M)\backslash\{0\}} \frac{\|\nabla f\|^2}{\|f\|^2}. \tag{9.81}$$

We will prove Theorem 9.39 using a unitary group strategy due to Chernoff [21]. A *unitary group* is a family of unitary operators $U(t)$, parametrized by $t \in \mathbb{R}$, such that

$$U(s)U(t) = U(s+t).$$

This concept was introduced in Exercise 5.5, where we saw that a self-adjoint operator A generates a unitary group $U(t) := e^{itA}$.

In the current application, we will use the solution of the wave equation to construct a unitary group whose generator is the Laplacian. The following result will then allow us to prove essential self-adjointness.

Theorem 9.40. *Suppose that A is a symmetric operator on a Hilbert space $\mathcal{H}$, such that A maps $\mathcal{D}(A)$ into itself. Let $U(t)$ be a unitary group on $\mathcal{H}$ with the following properties:*

(i) $U(t)A = AU(t)$ *on* $\mathcal{D}(A)$.
(ii) $U(t)\mathcal{D}(A) \subset \mathcal{D}(A)$.
(iii) *For all* $v \in \mathcal{D}(A)$,

$$\lim_{h \to 0} \frac{1}{h}\Big[U(t+h)v - U(t)v\Big] = iAU(t)v.$$

Then A^k is essentially self-adjoint on $\mathcal{D}(A)$ for each $k \in \mathbb{N}$.

Proof By Theorem 3.30, it suffices to establish the injectivity of $(A^k)^* \pm i$. Suppose that $(A^k)^* w = \pm i w$ for $w \in \mathcal{D}(A^*)$. For $v \in \mathcal{D}(A)$, consider the function

$$f(t) := \langle w, U(t)v \rangle. \tag{9.82}$$

By the assumption on the derivative of $U(t)$,

$$\begin{aligned} f^{(k)}(t) &= \Big\langle w, (iA)^k U(t) v \Big\rangle \\ &= i^k \Big\langle (A^k)^* w, U(t) \Big\rangle \\ &= \mp i^{k+1} f(t). \end{aligned}$$

The possible solutions are linear combinations of functions $e^{\alpha_j t}$, where $\{\alpha_j\}$ are the kth roots of $\mp i^{k+1}$. Since all of the roots have nonzero real part, any nontrivial solution will grow exponentially in at least one direction. On the other hand, $f(t)$ is bounded for all $t \in \mathbb{R}$ by the definition (9.82), because $U(t)$ is unitary. Therefore, $f \equiv 0$.

This shows in particular that $\langle w, v \rangle = 0$ for all $v \in \mathcal{D}(A)$, which implies $w = 0$ by the density of $\mathcal{D}(A)$. In other words, we have proven that

$$\ker\big((A^k)^* w \pm i\big) = \{0\}.$$

Therefore, A^k is essentially self-adjoint by Theorem 3.30. □

The wave group on a complete manifold can be constructed from wave solutions on a compact manifold, by virtue of the finite propagation speed established in Section 9.6.1.

Theorem 9.41. *Let M be a complete Riemannian manifold. Given initial data $u_0, v_0 \in C_0^\infty(M)$, the wave equation,*

$$\partial_t^2 u - \Delta u = 0, \qquad u(0, \cdot) = u_0, \qquad \partial_t u(0, \cdot) = v_0,$$

has a unique solution $u \in C^\infty(\mathbb{R} \times M)$ such that $u(t, \cdot)$ has compact support for each t.

Proof Since the compact case is covered by Theorem 9.37, we will assume that M is non-compact. Let $K_0 \subset M$ be a compact set containing the supports of u_0 and v_0. By the Heine–Borel property, the set

$$K_j := \{p \in M : \operatorname{dist}(p, K) \le j\}$$

is compact for $j \in \mathbb{N}$. For each j, we can define a compact manifold with boundary $\overline{\Omega}_j \subset M$ by smoothing the boundary of $K_{j+\frac{1}{2}}$, if necessary, so that $K_j \subset \Omega_j$.

Let $u_j(x, t)$ be the solution of the wave equation with Dirichlet boundary conditions on $\overline{\Omega}_j$, as given by Theorem 9.37, with initial data u_0, v_0. By Theorem 9.38, if $i < j$, then the solution $u_j(\cdot, t)$ has support within K_i for $|t| \le i$. Therefore, the restriction of u_j to $\overline{\Omega}_i$ is a solution of the wave equation on $\overline{\Omega}_i$ for $|t| \le i$. By the uniqueness property proven in Theorem 9.37, $u_j(x, t) = u_i(x, t)$ for $x \in \overline{\Omega}_i$ and $|t| \le i$. We can therefore patch together a global solution to the wave equation on M by setting

$$u(t, x) := \begin{cases} u_j(t, x), & |t| \le j \text{ and } x \in \Omega_j, \\ 0, & \text{otherwise.} \end{cases}$$

Uniqueness of the global solution follows from the conservation of energy argument used in Theorem 9.37. □

To obtain a unitary group from Theorem 9.41, we first introduce a new Hilbert space whose norm is motivated by conservation of energy. Recall the energy functional defined in (9.76),

$$E := \frac{1}{2}\|\partial_t u\|^2 + \frac{1}{2}\|\nabla u\|^2. \tag{9.83}$$

Let us define $\mathcal{W}$ as the Hilbert space completion (as described in Section 2.4) of $C_0^\infty(M)$ with respect to the norm $f \mapsto \|\nabla f\|$. The square of the norm of $(u, \partial_t u)$, as an element of $\mathcal{W} \oplus L^2(M)$, is then equal to $2E$.

On $C_0^\infty(M) \oplus C_0^\infty(M)$, we introduce the one-parameter family of maps

$$U(t)(u_0, v_0) = (u(\cdot, t), \partial_t u(\cdot, t)) \tag{9.84}$$

for $t \in \mathbb{R}$, where u is the wave solution associated with (u_0, v_0) by Theorem 9.41. The fact that E is constant for solutions of the wave equation implies that $U(t)$ is an isometry with respect to the norm on $\mathcal{W} \otimes L^2(M)$.

Clearly $U(0) = I$, and the uniqueness of wave solutions implies that $U(t_1+t_2) = U(t_1)U(t_2)$. Thus $U(t)$ is invertible on $C_0^\infty(M) \oplus C_0^\infty(M)$, with $U(t)^{-1} = U(-t)$. Since $C_0^\infty(M) \oplus C_0^\infty(M)$ is dense in $\mathcal{W} \oplus L^2(M)$, we obtain the following:

Theorem 9.42. *The family of operators $U(t)$ defined by (9.84) extends to a unitary group action on $\mathcal{W} \oplus L^2(M)$.*

This result gives in particular the existence of weak solutions of the wave equation on M, given initial data in $\mathcal{W} \otimes L^2(M)$. It also leads to a proof of Theorem 9.39, through the following:

Proof of Theorem 9.39 For the unitary group $U(t)$ given by $\mathcal{W} \oplus L^2(M)$, let $\mathcal{D}(A) = C_0^\infty(M) \oplus C_0^\infty(M)$. By construction, for $u_0, v_0 \in \mathcal{D}(A)$, we have

$$\lim_{h\to 0} \frac{1}{h}\Big[U(t+h) - U(t)\Big](u_0, v_0) = \partial_t(u, \partial_t u)$$
$$= (\partial_t u, \Delta u).$$

Hence, the requirements of Theorem 9.40 are satisfied with

$$A := \begin{pmatrix} 0 & -iI \\ -i\Delta & 0 \end{pmatrix}.$$

Therefore, the operator

$$A^2 = \begin{pmatrix} -\Delta & 0 \\ 0 & -\Delta \end{pmatrix}$$

is essentially self-adjoint on $\mathcal{D}(A)$. It follows that $-\Delta$ is essentially self-adjoint on $C_0^\infty(M)$. □

To apply the strategy used in the proof of Theorem 9.39 to a complete manifold with boundary $\overline{\Omega}$, we need a different choice of core domain. This is because the space $C_0^\infty(\overline{\Omega})$ is not necessarily preserved under the wave group, even when $\partial\Omega$ is smooth. Since the issue here is regularity, we can resolve it by imposing weaker regularity assumptions on the core. For example, if we take

$$\mathcal{D}_0 := \big\{u \in L^2(\Omega) : \ u \in \mathcal{D}(-\Delta_K) \text{ for some compact } K \subset \overline{\Omega}\big\}, \tag{9.85}$$

where $-\Delta_K$ denotes the Dirichlet Laplacian on K, then Theorem 9.38 shows that $\mathcal{D}_0$ is preserved under the wave group. We can therefore argue as above to deduce that $-\Delta$ is essentially self-adjoint on $\mathcal{D}_0$.

The larger core domain (9.85) has the advantage of not requiring any regularity of $\partial\Omega$. We can use it, for example, to construct a unique self-adjoint extension of the Laplacian on $M\backslash K$ where M is a complete Riemannian manifold and K is compact.

9.8 Essential Spectrum of Complete Manifolds

For a complete Riemannian manifold, the bottom of the spectrum and bottom of the essential spectrum are of great interest from both physical and geometric points of view. These are both geometric invariants, by the uniqueness of the self-adjoint extension described in Theorem 9.39. Of course, for a compact manifold the bottom of the spectrum is zero and the essential spectrum is empty, so these quantities are only interesting in the non-compact case.

In this section, we will prove that the bottom of the essential spectrum depends only on the geometry "at infinity," and then use this to establish an estimate based on volume growth.

9.8.1 Decomposition Principle

In Corollary 7.13, we saw that a bounded, compactly supported perturbation of a potential does not affect the essential spectrum of a Schrödinger operator. It is thus natural to expect that compactly supported perturbations of a complete metric would not change the essential spectrum. In fact, we can go even farther and consider perturbations that change the structure of the manifold within a compact set.

Let M be a complete Riemannian manifold with a compact subset $K \subset M$. Let $-\Delta_M$ denote the Laplacian on M, and $-\Delta_{M\setminus K}$ the self-adjoint Laplacian on $M\setminus K$ defined by imposing Dirichlet boundary conditions on ∂K, as described at the end of Section 9.7. We will prove the following result, due to Donnelly and Li [27, Prop. 2.1], which implies that a pair of manifolds which are isometric outside a compact set have the same essential spectrum.

Theorem 9.43. *Let M be a complete Riemannian manifolds. If $K \subset M$ is compact, then*

$$\sigma_{\mathrm{ess}}(-\Delta_M) = \sigma_{\mathrm{ess}}(-\Delta_{M\setminus K}).$$

Proof Recall that for a self-adjoint operator, the essential spectrum was characterized in Theorem 5.13 by the existence of a Weyl sequence. The strategy for the proof is to start with a Weyl sequence for $\lambda \in \sigma_{\mathrm{ess}}(-\Delta_M)$, and then construct a new Weyl sequence consisting of functions supported in $M\setminus K$.

As noted after the proof of Theorem 5.13, it suffices to consider Weyl sequences contained in a core domain on which the operator is essentially self-adjoint. Thus, for $\lambda \in \sigma_{\mathrm{ess}}(-\Delta_M)$ there exists a sequence $\{\varphi_k\} \subset C_0^\infty(M)$ such that $\|\varphi_k\| = 1$, $\varphi_k \to 0$ weakly, and

$$(-\Delta - \lambda)\varphi_k \to 0 \tag{9.86}$$

in norm. The sequence is also bounded uniformly in the H^1 norm, by (9.86) and the fact that

$$\|\varphi_k\|_{H^1}^2 = \big\langle \varphi_k, (-\Delta + 1)\varphi_k \big\rangle,$$

from Green's identity.

Choose a pair of cutoff functions $\chi, \chi_1 \in C_0^\infty(M)$ such that $\chi = 1$ on K, and $\chi_1 = 1$ on $\operatorname{supp} \chi$. Elliptic regularity (Theorem 9.27) gives the estimate

$$\|\chi_1\varphi_k\|_{H^2} \le C\big(\|\Delta(\chi_1\varphi_k)\| + \|\chi_1\varphi_k\|\big). \tag{9.87}$$

Note that

$$\Delta(\chi_1\varphi_k) = \chi_1(\Delta + \lambda)\varphi_k - \lambda\chi_1\varphi_k + [\Delta, \chi_1]\varphi_k.$$

The first term on the right approaches zero as $k \to \infty$ by (9.86), and the second and third terms are uniformly bounded by the H^1 bounds on $\{\varphi_k\}$. Therefore, by (9.87), the H^2 norms of $\{\chi_1\varphi_k\}$ are uniformly bounded in k.

Let $\overline{\Omega} \subset M$ be a compact manifold with boundary containing the support of χ_1. By the uniform H^2 bound on $\{\chi_1\varphi_k\}$, the sequences $\{\nabla(\chi_1\varphi_k)\}$ and $\{\chi_1\varphi_k\}$ are bounded in $H_0^1(\Omega)$. We can therefore apply Rellich (Theorem 6.9) in a finite collection of coordinate patches. By passing to a subsequence, we can therefore assume that $\{\chi_1\varphi_k\}$ converges in $H^1(\Omega)$. Since $\varphi_k \to 0$ weakly, the H^1 convergence implies that

$$\lim_{k\to\infty} \|\chi_1\varphi_k\|_{H^1} \to 0. \tag{9.88}$$

To complete the proof, consider the normalized sequence

$$\psi_k := \frac{(1-\chi)\varphi_k}{\|(1-\chi)\varphi_k\|},$$

which is supported in $M\backslash K$. We claim that this is a Weyl sequence for λ. The normalization factor in the denominator satisfies

$$\lim_{k\to\infty} \|(1-\chi)\varphi_k\| = 1,$$

by (9.88) and the fact that $\|\varphi_k\| = 1$. It follows that $\psi_k \to 0$ weakly, by the weak convergence $\varphi_k \to 0$.

To complete the proof, we check the convergence of

$$(-\Delta - \lambda)\psi_k = \frac{1}{\|(1-\chi)\varphi_k\|}\Big((1-\chi)(-\Delta-\lambda)\varphi_k + [\Delta, \chi]\varphi_k\Big).$$

The normalizing factor approaches one, as noted above. As $k \to \infty$, the first term in parentheses on the right converges to zero by (9.86), while the second term converges to zero by (9.88). Therefore,

$$(-\Delta - \lambda)\psi_k \to 0,$$

establishing that $\{\psi_k\}$ is a Weyl sequence for λ. □

9.8.2 The Bottom of the Essential Spectrum

For a complete manifold, we noted that the bottom of the spectrum can be estimated by the Rayleigh quotient (9.81). This formulation extends to the Laplacian on $M\backslash K$, where $K \subset M$ is compact and $-\Delta_{M\backslash K}$ denotes the Dirichlet Laplacian as in Section 9.8.1:

$$\inf \sigma(-\Delta_{M\backslash K}) = \inf_{f \in C_0^\infty(M\backslash K)} \frac{\|\nabla f\|^2}{\|f\|^2}. \tag{9.89}$$

This follows from the fact that $C_0^\infty(M\backslash K)$ is dense in the core domain (9.85) used to define $-\Delta_{M\backslash K}$.

The formula (9.89) gives an effective tool for locating the bottom of the full spectrum. The following result will allow us to apply similar estimates to the bottom of the essential spectrum.

Theorem 9.44. *For a complete manifold M,*

$$\inf \sigma_{\mathrm{ess}}(-\Delta_M) = \sup_{K \subset M} \Big(\inf \sigma(-\Delta_{M\backslash K})\Big),$$

where K ranges over all compact subsets of M.

Proof Let $a_0 := \inf \sigma_{\mathrm{ess}}(-\Delta_M)$. From Theorem 9.43 we immediately have

$$\sup_{K \subset M} \Big(\inf \sigma(-\Delta_{M\backslash K})\Big) \le a_0. \tag{9.90}$$

Assume, for the sake of contradiction, the inequality is strict. Then there exists $\varepsilon > 0$ such that

$$\inf \sigma(-\Delta_{M\backslash K}) \le a_0 - 2\varepsilon \tag{9.91}$$

for all compact $K \subset M$. By (9.89), for each K we can choose $\varphi \in C_0^\infty(M\backslash K)$ such that $\|\varphi\| = 1$ and

$$\langle \varphi, -\Delta\varphi \rangle < a_0 - \varepsilon. \tag{9.92}$$

Start by choosing such a function φ_1 for $K = \emptyset$. Then set $K_1 = \operatorname{supp} \varphi_1$ and choose $\varphi_2 \in C_0^\infty(M\backslash K_1)$. Continuing this process inductively yields an infinite orthonormal sequence of functions $\varphi_k \in C_0^\infty(M)$, each satisfying (9.92).

Let Π denote the spectral resolution of $-\Delta_M$. Since a_0 is the bottom of the essential spectrum, $\Pi_{[0,a_0-\varepsilon)}$ has finite rank. Therefore, there exists a nonzero function

$$u \in \operatorname{span}\{\varphi_k\} \cap \operatorname{range} \Pi_{[a_0-\varepsilon,\infty)}.$$

By (9.92), we have

$$\langle u, (-\Delta - a_0 + \varepsilon)u \rangle < 0.$$

On the other hand, the fact that $u \in \operatorname{range} \Pi_{[a_0-\varepsilon,\infty)}$ implies

$$\langle u, -\Delta u \rangle \ge (a_0 - \varepsilon)\|u\|^2.$$

This contradiction shows that such a function u cannot exist. Therefore the inequality (9.91) is impossible, which proves that equality holds in (9.90). □

By combining (9.89) with Theorem 9.44, we obtain a useful variational characterization of the bottom of the essential spectrum:

$$\inf \sigma_{\mathrm{ess}}(-\Delta_M) = \sup_{K \subset M} \left(\inf_{f \in C_0^\infty(M \setminus K)} \frac{\|\nabla f\|^2}{\|f\|^2} \right). \tag{9.93}$$

9.8.3 Volume Growth Estimate

As an example of the application of Theorem 9.44, in this section we will develop a simple bound on the essential spectrum based on volume growth. This result is adapted from Brooks [15].

Let M be a complete manifold. Fix an arbitrary base point $x_0 \in M$, and let

$$V(r) := \mathrm{vol}(B(x_0; r)).$$

The exponential rate of volume growth for M is defined as

$$\mu := \limsup_{r \to \infty} \frac{\log V(r)}{r}, \tag{9.94}$$

A simple argument with the triangle inequality shows that μ is independent of x_0. The Euclidean plane has $V(r) = \omega_n r^n$, so $\mu = 0$ in that case.

Our main goal is the following:

Theorem 9.45. *For a complete manifold M of infinite volume,*

$$\inf \sigma_{\mathrm{ess}}(-\Delta_M) \le \frac{\mu^2}{4}.$$

Example 9.46. Consider the Poincaré disk $\mathbb{B}$ introduced in Example 9.14. In geodesic polar coordinates (r, θ), the hyperbolic metric (9.25) takes the form

$$ds^2 = dr^2 + \sinh^2 r \, d\theta^2.$$

Hence

$$\begin{aligned} V(r) &= 2\pi \int_0^r \sinh(t)\, dt \\ &= 2\pi(\cosh r - 1). \end{aligned}$$

The volume growth rate is therefore $\mu = 1$, and Theorem 9.45 gives $\inf \sigma_{\mathrm{ess}}(-\Delta_{\mathbb{B}}) \le \frac{1}{4}$. In fact, we will see in Exercise 9.2 that $\sigma(-\Delta_{\mathbb{B}}) = [\frac{1}{4}, \infty)$, so the estimate is sharp. ♢

The proof of Theorem 9.45 follows immediately from Theorem 9.44 and the following:

Theorem 9.47. *If M is a complete manifold with boundary that has growth constant μ, defined as in (9.94), then*

$$\inf \sigma(-\Delta_M) \le \frac{\mu^2}{4}.$$

If M has infinite volume, then the same estimate holds for the Dirichlet Laplacian $-\Delta_{M \setminus K}$ for $K \subset M$ compact.

Proof Both sides of the inequality are zero if M is compact, so we can assume that M is not compact. Fix a base point $x_0 \in M$, and let $\rho(x) := \mathrm{dist}(x, x_0)$ and $V(r) = \mathrm{vol}\, B(x_0; r)$. For $\alpha > 0$, we have a simple integral estimate

$$\begin{aligned} \int_M e^{-\alpha\rho}\, dV &\le \sum_{m=0}^{\infty} \big[V(m+1) - V(m)\big] e^{-\alpha m} \\ &= \sum_{m=1}^{\infty} V(m) e^{-\alpha m}(e^{\alpha} - 1). \end{aligned}$$

By the definition of μ, we thus have

$$\int_M e^{-\alpha\rho}\, dV < \infty \tag{9.95}$$

for all $\alpha > \mu$.

The strategy is to exploit (9.95) by constructing a family of trial functions for the Rayleigh quotient. We first need to consider the regularity of ρ, since this function is not differentiable in general. By the triangle inequality,

$$|\rho(x) - \rho(y)| \le \mathrm{dist}(x, y),$$

so ρ is at least Lipschitz continuous with constant 1. We claim that this implies that ρ is weakly differentiable with $|\nabla\rho|_g \le 1$. This essentially follows from Rademacher's theorem, which says that a Lipschitz function on $\mathbb{R}^n$ is differentiable almost everywhere. For the proof see, for example, Federer [30, Thm. 3.1.6] or Heinonen [43, Thm. 3.1]. Applying Rademacher's theorem in local coordinate patches shows that ρ is differentiable almost everywhere, with $|\nabla\rho|_g \le 1$ where the derivative exists. Since Lipschitz implies absolute continuity, we can use integration by parts along line segments to conclude that the almost everywhere defined classical derivatives are also weak derivatives.

For $m \geq 0$, choose $\chi_m \in C_0^\infty(M)$ with $0 \leq \chi_m \leq 1$, such that

$$\chi_M(x) = \begin{cases} 1, & \rho(x) \leq m, \\ 0, & \rho(x) \leq m+1. \end{cases}$$

Since the distance between $\{x : \rho(x) = m\}$ and $\{x : \rho(x) = m+1\}$ equals 1 for all m, we can assume a uniform bound $|\nabla \chi_m|_g \leq c$ for all m.

Consider now the trial function $f_m := \chi_m e^{-\alpha\rho/2}$. Because the weak first derivatives of ρ are bounded, as noted above, we have $f_m \in H^1(M)$. The fact that f_m has compact support thus implies that f_m can be approximated by $C_0^\infty(M)$ functions with respect to the H^1 norm. Therefore, from the Rayleigh quotient (9.81) we deduce that

$$\sigma_0 \|f_m\|^2 \leq \|\nabla f_m\|^2, \tag{9.96}$$

where $\sigma_0 := \inf \sigma(-\Delta)$.

The right-hand side of (9.96) can be estimated by

$$\begin{aligned} \|\nabla f_m\|^2 &= \int_M \left[|\nabla \chi_m|_g^2 - \alpha \chi_m g(\nabla \chi_m, \nabla \rho) + \frac{\alpha^2}{4} \chi_m^2 |\nabla \rho|_g^2 \right] e^{-\alpha\rho} \, dV \\ &\leq (c^2 + c\alpha) \int_{\operatorname{supp} \nabla \chi_m} e^{-\alpha\rho} \, dV + \frac{\alpha^2}{4} \int_M \chi_m^2 e^{-\alpha\rho} dV, \end{aligned}$$

using the gradient bounds $|\nabla \chi_m|_g \leq c$ and $|\nabla \rho|_g \leq 1$. For $\alpha > \mu$, $e^{-\alpha\rho}$ is integrable by (9.95), so the integral over $\operatorname{supp} \nabla \chi_m$ vanishes as $m \to \infty$, yielding

$$\limsup_{m \to \infty} \|\nabla f_m\|^2 \leq \frac{\alpha^2}{4} \|e^{-\alpha\rho}\|_{L^1}.$$

Since $\|f_m\|^2 \to \|e^{-\alpha\rho}\|_{L^1}$ as $m \to \infty$, taking the limit of (9.96) gives

$$\sigma_0 \leq \frac{\alpha^2}{4}$$

for all $\alpha > \mu$. This proves the claim for $-\Delta_M$.

For the case $M \backslash K$, we need the trial functions to vanish near K, which changes the argument somewhat. As above, fix a base point $x_0 \in M$ and let $\rho(x) := \operatorname{dist}(x, x_0)$. We also fix an inner cutoff function $\psi \in C_0^\infty(M)$, with $\psi = 1$ on K. For $k \geq 0$ set

$$\rho_k(x) := \begin{cases} 0, & \rho(x) \leq k, \\ \rho(x) - k, & \rho(x) > k. \end{cases}$$

The new family of trial functions is

$$f_{m,k} = \chi_m(1-\psi)e^{-\alpha\rho_k/2},$$

which by (9.89) satisfies

$$\sigma_0(M\backslash K)\left\|f_{m,k}\right\|^2 \le \left\|\nabla f_{m,k}\right\|^2.$$

Assuming that $\alpha > \mu$, we can estimate $\|\nabla f_{m,k}\|^2$ and take the limit $m \to \infty$ as above. Assuming that k is large enough that $\rho_k = 0$ on $\operatorname{supp}\psi$, this gives

$$\limsup_{m\to\infty}\left\|\nabla f_{m,k}\right\|^2 \le \frac{\alpha^2}{4}\left\|(1-\psi)^2 e^{-\alpha\rho_k}\right\|_{L^1} + \|\nabla\psi\|^2.$$

Thus, for $\alpha > \mu$ and k sufficiently large,

$$\left(\sigma_0(M\backslash K) - \frac{\alpha^2}{4}\right)\left\|(1-\psi)^2 e^{-\alpha\rho_k}\right\|_{L^1} \le \|\nabla\psi\|^2. \tag{9.97}$$

The fact that $\rho_k \to 0$ pointwise and the assumption that M has infinite volume together imply that

$$\lim_{k\to\infty}\left\|(1-\psi)^2 e^{-\alpha\rho_k}\right\|_{L^1} = \infty.$$

Therefore, (9.97) gives the inequality

$$\sigma_0(M\backslash K) \le \frac{\alpha^2}{4}$$

for all $\alpha > \mu$, which proves the claim. □

9.9 Exercises

9.1. Another model for two-dimensional hyperbolic space is the upper half-plane $\mathbb{H} := \{(x, y) : \ y > 0\}$, with the metric

$$ds^2 = \frac{dx^2 + dy^2}{y^2}.$$

(a) If $\mathbb{H}$ and $\mathbb{B}$ are viewed as subsets of $\mathbb{C}$, then the Cayley transform,

$$z \mapsto \frac{z-i}{z+i},$$

maps $\mathbb{H}$ to $\mathbb{B}$. Show that this map is an isometry.

(b) Compute the Laplacian $-\Delta_{\mathbb{H}}$ in the $\mathbb{H}$ model, and show that the function y^s satisfies the eigenvalue equation $-\Delta_{\mathbb{H}} u = \lambda u$ for $s \in \mathbb{C}$.
(c) Show that

$$[\tfrac{1}{4}, \infty) \subset \sigma(-\Delta_{\mathbb{H}}),$$

by creating a Weyl sequence using y^s with cutoff functions.

9.2. Let $-\Delta_{\mathbb{H}}$ denote the hyperbolic Laplacian for the upper half-space introduced in Exercise 9.1. In this exercise we will show that

$$-\Delta_{\mathbb{H}} - \tfrac{1}{4} \geq 0,$$

which shows that the bound from Example 9.46 is sharp.

(a) Use integration by parts and Cauchy–Schwarz to show that

$$\int_{-\infty}^{\infty} |\phi|^2 \frac{dy}{y^2} \leq 4 \int_{-\infty}^{\infty} \left| \frac{\partial \phi}{\partial y} \right|^2 dy,$$

with x fixed, for $\phi \in C_0^\infty(\mathbb{H})$.
(b) Deduce from (a) that

$$\int_{\mathbb{H}} \left(\left| \frac{\partial \phi}{\partial x} \right|^2 + \left| \frac{\partial \phi}{\partial y} \right|^2 \right) dx\, dy \geq \frac{1}{4} \int_{\mathbb{H}} |\phi|^2 \frac{dx\, dy}{y^2}$$

for $\phi \in C_0^\infty(\mathbb{H})$. Show that this implies the claimed lower bound on $-\Delta_{\mathbb{H}}$.

9.3. Let $\mathbb{S}^n$ denote the unit sphere in $\mathbb{R}^{n+1}$. By writing the Euclidean metric in geodesic polar coordinates, we can see that

$$-\Delta_{\mathbb{R}^{n+1}} = -\frac{1}{r^n} \frac{\partial}{\partial r} \left(r^n \frac{\partial}{\partial r} \right) - \frac{1}{r^2} \Delta_{\mathbb{S}^n}.$$

Let $q(x)$ be a harmonic polynomial on $\mathbb{R}^{n+1}$, which means that

$$\Delta_{\mathbb{R}^{n+1}} q = 0.$$

Assume also that q is homogeneous of degree $l = 0, 1, 2, \ldots$, meaning that

$$q(cx) = c^l q(x)$$

for some $c > 0$.

(a) Show that the restriction of a homogeneous harmonic polynomial q to $\mathbb{S}^n$ is an eigenfunction of $-\Delta_{\mathbb{S}^n}$, and compute the eigenvalue. (This construction yields the full spectrum of $-\Delta_{\mathbb{S}^n}$.)
(b) In the case $n = 2$, show that the spherical harmonics Y^l_m discussed in Section 7.4 are restrictions to the unit sphere of harmonic polynomials on $\mathbb{R}^3$.

9.4. In a compact Riemannian manifold M, let $B(x; r)$ denote a metric ball of radius r centered at x, not necessarily contained in a geodesic coordinate neighborhood. Suppose that we have estimates for the first Dirichlet eigenvalue of a ball, of the form

$$\lambda_1(B(x; r)) \leq \beta(r)$$

for all $x \in M$ and $r > 0$, where β is a function on $\mathbb{R}^+$. (Estimates of this type can be deduced from lower bounds on the curvature.) Use the min–max principle to prove that

$$\lambda_k(M) \leq \beta(d/2k),$$

where $d := \operatorname{diam}(M)$.

9.5. For a compact manifold M, prove that there exists a constant $C > 0$ such that if ϕ is a normalized eigenfunction of $-\Delta$ with eigenvalue λ, then

$$\sup |\phi| \leq C\lambda^{\frac{n-1}{4}}.$$

9.6. Let M be a compact Riemannian manifold. The *scalar curvature* S at a point $x \in M$ can be defined as a coefficient in the expansion for the area of a small geodesic sphere. That is, $S(x)$ is defined by the fact that

$$\operatorname{area}(\partial B_M(x; \varepsilon)) = \operatorname{area}(\partial B_{\mathbb{R}^n}(x; \varepsilon))\left[1 + \frac{S(x)}{6n}\varepsilon^2 + O(\varepsilon^4)\right]$$

as $\varepsilon \to 0$, where B_M denotes a metric ball in M and $B_{\mathbb{R}^n}$ a ball in Euclidean $\mathbb{R}^n$. Compute the coefficient $u_1(x, x)$ from heat parametrix formula (9.66) in terms of $S(x)$. Show that the resulting heat coefficient in Theorem 9.34 is

$$a_1 = \frac{1}{6}\int_M S\, dV.$$

Notes

For a thorough introduction to differentiable manifolds, see Lee [58]. The background material on differential geometry presented in Section 9.2 is covered

much more completely in the standard texts. See, for example, do Carmo [26], Klingenberg [52], Lee [57], or Petersen [66].

The spectrum of a compact Riemannian manifold is discussed in greater detail in Berger [8, Ch. 9], Berger, Gauduchon, and Mazet [9], Chavel [19], and Schoen and Yau [81, Ch. III]. Much of the focus in these books is on estimates of eigenvalues in terms of Riemannian curvature, which we did not get into here.

The spectral theory of hyperbolic manifolds is covered in particular in Borthwick [14] and Buser [17]. For applications to automorphic forms and number theory, see Iwaniec [47] or Venkov [92].

The issue of recovering geometric data from the spectrum is reviewed in the survey by Brooks [16]. See Lablée [55] for a recent more survey of spectral geometry.

For more details on the behavior of heat kernels in the Riemannian setting see Davies [23, Ch. 5] or Schoen and Yau [81, Ch. IV].

Appendix A
Background Material

A.1 Measure and Integration

As noted in the text, a *measure* on a set X consists of a σ-algebra $\mathcal{M}$ of measurable subsets of X and a countably additive function $\mu : \mathcal{M} \to [0, \infty]$. To recall the basic definitions, a σ-algebra is a collection of subsets that is closed under countable unions and complements (and hence countable intersections as well). The countable additivity property means that for any countable disjoint sequence of sets $A_1, A_2, \ldots,$

$$\mu\left(\cup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j). \tag{A.1}$$

In the text we considered only σ-finite measure spaces, for which X can be decomposed into a countable union of sets of finite measure.

Two important special cases can be defined with $\mathcal{M}$ equal to the collection of all subsets of X:

Example A.1. The *counting measure* ν is defined by

$$\nu(A) := \begin{cases} \#A, & A \text{ is finite,} \\ \infty, & A \text{ is infinite.} \end{cases}$$

Counting measure is the default choice for a discrete set, such as $\mathbb{N}$ or $\mathbb{Z}$. (Clearly, ν is σ-finite if and only if X is countable.)

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1

Example A.2. For $p \in X$, the *point measure* δ_p is defined by

$$\delta_p(A) := \begin{cases} 1, & p \in A, \\ 0, & p \notin A. \end{cases}$$

For a topological space X, the collection $\mathcal{B}$ of *Borel sets* is defined as the σ-algebra generated by the open subsets of X. A *Borel measure* is simply a measure defined on $\mathcal{B}$.

In this section we will review some of the basics of measure and integration theory. This material is standard and covered in many texts, so we will omit most of the proofs. For additional background, see, e.g., Folland [31], Royden [76], Rudin [78], or Stein and Shakarchi [87].

A.1.1 Lebesgue Measure

To define Lebesgue measure on $\mathbb{R}^n$, we start from the standard definition of the volume of a closed rectangle,

$$\text{vol}\big(I_1 \times I_2 \times \cdots \times I_n\big) := \prod_{j=1}^{n} \ell(I_j),$$

where $\ell[a,b] := b - a$. Let $\mathcal{R}$ denote the collection of closed rectangles in $\mathbb{R}^n$. The *outer measure* of a set $A \subset \mathbb{R}^n$ is then defined by taking the infimum over coverings by countable unions of rectangles,

$$m^*(A) := \inf\left\{ \sum_{j=1}^{\infty} \text{vol}(R_j) \,:\, A \subset \bigcup_{j=1}^{\infty} R_j \right\}.$$

In general, an outer measure is defined as a countably subadditive set function, meaning that

$$m^*(\cup A_j) \le \sum_{j=1}^{\infty} m^*(A_j),$$

for a countable sequence of sets $\{A_j\}$.

To obtain a measure, we need to restrict m^* to an appropriate class of measurable sets. Constantin Carathéodory established a criterion for this: a set E is Lebesgue measurable if, for each $A \subset \mathbb{R}^n$,

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c). \tag{A.2}$$

This condition defines the collection $\mathcal{M}$ of measurable sets, which forms a σ-algebra. Lebesgue measure is defined as the restriction,

$$m := m^*|_{\mathcal{M}}.$$

The criterion (A.2) can be used to produce a measure μ from any outer measure μ^*.

The Lebesgue class $\mathcal{M}$ includes all open sets. This is not immediately clear from the condition (A.2), but in fact one can show that $E \in \mathcal{M}$ if and only if for each $\varepsilon > 0$ there exists an open set $U \supset E$ such that

$$m^*(U \backslash E) \leq \varepsilon. \tag{A.3}$$

A similar characterization in terms of closed sets contained in E is also possible. Since $\mathcal{M}$ is a σ-algebra, the fact that it contains all open sets implies that $\mathcal{B} \subset \mathcal{M}$.

Since Lebesgue measure generalizes the classical notion of volume, it is common to denote $m(A)$ by $\mathrm{vol}(A)$, especially in geometric contexts.

A.1.2 Integration

On a measure space $(X, \mathcal{M}, \mu)$, a *simple* function φ is a finite linear combination of characteristic functions,

$$\varphi = \sum_{j=1}^{m} c_j \chi_{E_j},$$

where $c_j \in \mathbb{C}$, $E_j \in \mathcal{M}$, and $\mu(E_j) < \infty$ for each j. The integral of a simple function is defined by the obvious sum,

$$\int_X \varphi \, d\mu := \sum_{j=1}^{m} c_j \mu(E_j).$$

A function $f : X \to \mathbb{R}^n$ is *measurable* if the preimage of each Borel set is contained in $\mathcal{M}$. The measurability of f implies that there exists a sequence of simple functions $\{\varphi_j\}$ such that $\varphi_j \to f$ pointwise and $|\varphi_j| \to |f|$ monotonically. With such an approximation we can define

$$\int_X |f| \, d\mu := \lim_{j \to \infty} \int_X |\varphi_j| \, d\mu.$$

We say that f is *integrable* if the integral of $|f|$ is finite, in which case we can define

$$\int_X f \, d\mu := \lim_{j \to \infty} \int_X \varphi_j \, d\mu.$$

The integral is well defined independently of the approximating sequence, and linear in the sense that for two integrable functions,

$$\int_X (f+g)\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu.$$

It is also monotonic, meaning that

$$f \le g \Longrightarrow \int_X f\, d\mu \le \int_X g\, d\mu.$$

In the case of Lebesgue measure on $\mathbb{R}^n$, this integral construction generalizes the ordinary Riemann integral. In the main text we write the Lebesgue integral on $\mathbb{R}^n$ in the traditional notation, replacing dm by $d^n x$.

The integral construction sketched here yields a trio of very useful convergence theorems.

Theorem A.3 (Monotone Convergence Theorem). *Suppose that* $\{f_j\}$ *is a sequence of measurable functions with*

$$0 \le f_1 \le f_2 \le \dots.$$

Then,

$$\lim_{j\to\infty} \int_X f_j\, d\mu = \int_X \left(\lim_{j\to\infty} f_j \right) d\mu$$

(where both sides could be infinite).

Theorem A.4 (Fatou's Lemma). *If* $\{f_j\}$ *is a sequence of measurable functions with* $f_j \ge 0$*, then*

$$\int_X \left(\liminf_{j\to\infty} f_j \right) d\mu \le \liminf_{j\to\infty} \int_X f_j\, d\mu$$

(where both sides could be infinite).

Theorem A.5 (Dominated Convergence Theorem). *Suppose that* $\{f_j\}$ *is a sequence of measurable functions such that* $f_j \to f$ *pointwise. If there exists an integrable function* g *such that* $|f_j| \le g$ *for all* j*, then*

$$\lim_{j\to\infty} \int_X f_j\, d\mu = \int_X f\, d\mu.$$

A.1.3 Product Measure

Given two measure spaces $(X_1, \mathcal{M}_1, \mu_1)$ and $(X_2, \mathcal{M}_2, \mu_2)$, we can construct a measure on $X_1 \times X_2$ using a generalization of the approach outlined in Appendix A.1.1. We start by defining the measure of a "rectangular set" in the obvious way. For $A \in \mathcal{M}_1$ and $B \in \mathcal{M}_2$,

$$\pi_0(A \times B) := \mu_1(A)\mu_2(B). \tag{A.4}$$

We then use coverings by rectangles to construct an outer measure, for $E \subset X_1 \times X_2$,

$$\pi^*(E) := \inf\left\{\sum_{j=1}^{\infty} \pi_0(A_j \times B_j) : A \subset \bigcup_{j=1}^{\infty} (A_j \times B_j)\right\}.$$

Carathéodory's condition defines a class $\mathcal{M}$ of subsets of $X_1 \times X_2$ measurable with respect to π^*. The restriction of π^* to $\mathcal{M}$ then defines the product measure π, which is commonly written as $\pi = \mu_1 \times \mu_2$.

In the Lebesgue case, we could apply the product construction to obtain the measure on $\mathbb{R}^n$ as a product of measures on $\mathbb{R}$.

In principle, integrating with respect to a product measure on $X_1 \times X_2$ could give different results from an iterated integral defined by integrating separately over the original X_1 and X_2. However, we can avoid this issue under fairly general conditions.

Theorem A.6 (Fubini). *Suppose that $(X_1, \mathcal{M}_1, \mu_1)$ and $(X_2, \mathcal{M}_2, \mu_2)$ are σ-finite measure spaces, with the product space $(X_1 \times X_2, \mathcal{M}, \mu_1 \times \mu_2)$. If f is an integrable function on $X_1 \times X_2$, then the iterated integrals make sense in either order and*

$$\begin{aligned}\int_{X_1 \times X_2} f \, d(\mu_1 \times \mu_2) &= \int_{X_1} \left(\int_{X_2} f(x, y) \, d\mu_2(y) \right) d\mu_1(x) \\ &= \int_{X_2} \left(\int_{X_1} f(x, y) \, d\mu_1(x) \right) d\mu_2(y).\end{aligned}$$

The same conclusion holds without the integrability assumption if $f \geq 0$.

A.1.4 Differentiation

A function $f : [a, b] \to \mathbb{C}$ is *absolutely continuous* if for each $\varepsilon > 0$ there exists $\delta > 0$ such that for every finite collection of disjoint subintervals (a_j, b_j) satisfying

$$\sum_{j=1}^{k} (b_j - a_j) < \delta,$$

we have

$$\sum_{j=1}^{k} \left| f(b_j) - f(a_j) \right| < \varepsilon.$$

One way to obtain an absolutely continuous function is by integration. If $h \in L^1[a, b]$, then the function

$$f(x) := \int_a^x h(t)\, dt \tag{A.5}$$

is absolutely continuous. Indeed, by a general measure theory argument, if $g \in L^1(X, d\mu)$, then for $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\mu(E) < \delta \quad \Longrightarrow \quad \int_E |g|\, d\mu < \varepsilon. \tag{A.6}$$

Applying this in the case of Lebesgue measure on $\mathbb{R}$, with E a finite union of intervals, shows that (A.5) is absolutely continuous.

It turns that all absolutely continuous functions can be expressed as definite integrals.

Theorem A.7 (Lebesgue Differentiation Theorem). *If f is absolutely continuous on $[a, b]$, then f' exists almost everywhere, $f' \in L^1[a, b]$, and*

$$f(x) = f(a) + \int_a^x f'(t)\, dt.$$

Conversely, for $g \in L^1(a, b)$ the function defined by

$$f(x) := \int_a^x g\, dt$$

is absolutely continuous, with $f' = g$ a.e.

The property (A.6) suggests a related definition for measures. On a measure space $(X, \mathcal{M}, \mu)$, a measure ν on $\mathcal{M}$ is absolutely continuous with respect to μ if for $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\mu(E) < \delta \quad \Longrightarrow \quad \nu(E) < \varepsilon.$$

By standard measure theory arguments, absolute continuity holds if and only if every set of measure zero with respect to μ also has measure zero with respect to ν. (This latter condition is frequently taken as the definition.)

A.1.5 Decomposition of Measures

In this section we will review the Lebesgue decomposition theorem on $\mathbb{R}$, which was applied in Section 5.4.2 to establish a classification of the spectrum.

A Borel measure on $\mathbb{R}$ is *regular* if $\mu(K) < \infty$ when K is compact. Let m denote the Lebesgue measure on $\mathbb{R}$. For the decomposition theorem, we distinguish the following types of Borel measure:

(i) A *pure point* measure is a linear combination of point measures.
(ii) A measure μ is *absolutely continuous* if $\mu(E) = 0$ whenever $m(E) = 0$.
(iii) A measure μ is *singular continuous* if μ is supported on a set of Lebesgue measure zero, but $\mu\{x\} = 0$ for each $x \in \mathbb{R}$.

The Cantor measure (a probability measure supported on the Cantor set) is the classic example of a singular continuous measure.

Theorem A.8 (Lebesgue Decomposition Theorem). *A regular Borel measure μ on $\mathbb{R}$ admits a unique decomposition,*

$$\mu = \mu_{\mathrm{pp}} + \mu_{\mathrm{ac}} + \mu_{\mathrm{sc}},$$

where μ_{pp} is pure point, μ_{ac} is absolutely continuous, and μ_{sc} is singular continuous.

Proof Consider the subset

$$Z := \{x \in \mathbb{R} : \ \mu\{x\} > 0\}.$$

Since $\mu[-n, n] < \infty$ for all n by the regularity assumption, $Z \cap [-n, n]$ is finite. Hence Z is at most countable. Let $Z = \{z_1, \ldots\}$, and define

$$\mu_{\mathrm{pp}} := \sum_j \mu\{z_j\}\delta_{z_j}.$$

If we then define $\mu_{\mathrm{c}} := \mu - \mu_{\mathrm{pp}}$, then μ_{c} is a continuous Borel measure, meaning that single point has measure zero. Let α be the associated cumulative distribution function, centered at 0,

$$\alpha(x) := \begin{cases} -\mu_{\mathrm{c}}(x, 0], & x < 0, \\ \mu_{\mathrm{c}}[0, x], & x \geq 0. \end{cases}$$

Since μ_{c} is a continuous measure, α is a continuous increasing function. It follows that α' exists almost everywhere (with respect to Lebesgue measure) and is locally integrable. We can thus define

$$\mu_{\rm ac}(E) := \int_E \alpha' \, dm,$$

and then set $\mu_{\rm sc} = \mu_{\rm c} - \mu_{\rm ac}$. □

For additional details on this construction, we refer the reader to Folland [31, §3.5].

A.1.6 Riesz Representation

Let X be a compact topological space. In this section, we will develop the version of the Riesz representation theorem stated as Theorem 5.3, which relates certain continuous linear functionals on $C(X)$ with Borel measures on X. Here $C(X)$ denotes the space of continuous functions $X \to \mathbb{C}$.

A linear functional $\beta : C(X) \to \mathbb{C}$ is *positive* if

$$\beta(f) \geq 0 \text{ for } f \geq 0.$$

Applying the positivity condition to $(\sup|f| - |f|) \geq 0$ shows that

$$\beta(|f|) \leq \beta(1) \sup|f|.$$

It follows that a positive functional is bounded with respect to the sup norm, because

$$\begin{aligned} |\beta(f)| &= \sqrt{\beta(\mathrm{Re}\, f)^2 + \beta(\mathrm{Im}\, f)^2} \\ &\leq \sqrt{2}\beta(1) \sup|f|. \end{aligned}$$

The existence of a partition of unity will play an important role in the proof. This is easy to establish for a compact metric space.

Lemma A.9 (Partition of Unity). *Let X be a compact metric space, and $\{U_j\}_{j=1}^m$ a finite open cover. There exists a set of functions $\psi_j \in C(X)$ such that $0 \leq \psi_j \leq 1$, $\operatorname{supp} \psi_j \subset U_j$ for each $j = 1, \ldots, n$, and*

$$\sum_{j=1}^m \psi_j = 1.$$

Proof For each $x \in X$, there exists an open metric ball B_x for which $\overline{B_x} \subset U_j$ for some j. Since X is compact, a finite number of these balls, say B_{x_i} for $i = 1, \ldots m$, is sufficient to cover X. Set $g_i(x) := \operatorname{dist}(x, B_{x_i}^c)$ so that g_i is continuous and $g_i(x) > 0$ if and only if $x \in B_{x_i}$.

For each j let h_j be the sum of the g_i for which $\overline{B_{x_i}} \subset U_j$. Since the B_{x_i} form a cover, for every $x \in X$ we have $h_j(x) > 0$ for at least one j. The desired functions ψ_j are then obtained by normalizing

$$\psi_j(x) := \frac{h_j(x)}{\sum_k h_k(x)}.$$

□

For $f \in C(X)$ and $U \subset X$ an open set, we use the notation

$$f \prec U$$

to mean that $0 \le f \le 1$ and $\operatorname{supp} f \subset U$.

Theorem A.10 (Riesz Representation Theorem). *Let X be a compact metric space. Given a positive linear functional $\beta : C(X) \to \mathbb{C}$, there exists a unique Borel measure μ on X such that*

$$\beta(f) = \int_X f\, d\mu \tag{A.7}$$

for $f \in C(X)$.

Proof For an open set $U \subset X$, let

$$\mu(U) := \sup\bigl\{\beta(f) : f \in C(X), f \prec U\bigr\}.$$

For an arbitrary subset $A \subset X$, we then set

$$\mu^*(A) := \inf\bigl\{\mu(U) : U \text{ open and } A \subset U\bigr\}. \tag{A.8}$$

To see that μ^* is an outer measure, we need to establish countable subadditivity. Suppose that $\{A_j\}$ is a countable sequence of subsets of X, and set $A := \cup A_j$. Our goal is to prove that

$$\mu^*(A) \le \sum \mu^*(A_j). \tag{A.9}$$

Given $\varepsilon > 0$, for each j we can find an open set $U_j \supset A_j$ such that

$$\mu(U_j) \le \mu^*(A_j) + 2^{-j}\varepsilon.$$

Adding these together thus gives

$$\sum_{j=1}^{\infty} \mu(U_j) \le \sum_{j=1}^{\infty} \mu^*(A_j) + \varepsilon. \tag{A.10}$$

Now let $U := \cup_j U_j$. By the definition (A.8),

$$\mu^*(A) \leq \mu^*(U). \tag{A.11}$$

Consider $f \in C(X)$ with $f \prec U$. Since $\operatorname{supp}(f)$ is compact, we have $\operatorname{supp}(f) \subset \cup_{j=1}^n U_j$ for some finite n. If we set $U_0 := \operatorname{supp}(f)^c$, then $\{U_j\}_{j=0}^\infty$ is a cover for X. By Lemma A.9, there exists a partition of unity $\{\psi_j\}_{j=0}^n \subset C(X)$, with $0 \leq \psi_j \leq 1$, $\operatorname{supp} \psi_j \in U_j$ and

$$\sum_{j=0}^n \psi_j = 1.$$

This construction yields $f = \sum_{j=1}^n f\psi_j$ and $\psi_j f \prec U_j$ for $j = 1, \ldots, n$. Hence,

$$\begin{aligned}\beta(f) &= \sum_{j=1}^n \beta(f\psi_j) \\ &\leq \sum_{j=1}^n \mu(U_j).\end{aligned}$$

By (A.10), this gives

$$\beta(f) \leq \sum_{j=1}^\infty \mu^*(A_j) + \varepsilon.$$

Since this holds for all $f \prec U$, it implies that

$$\mu^*(U) \leq \sum_{j=1}^\infty \mu^*(A_j) + \varepsilon.$$

Applying (A.11) and taking $\varepsilon \to 0$ thus prove (A.9).

With μ^* established as an outer measure, we can now apply the standard Carathéodory construction to obtain a measure μ on the σ-algebra $\mathcal{M}$ defined by the condition that $E \subset X$ is measurable if

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A - E) \tag{A.12}$$

for all $A \subset X$.

To show that μ is a Borel measure, we must check that each open set $U \subset X$ satisfies (A.12). Given $A \subset X$ and $\varepsilon > 0$, choose an open set $V \supset A$ such that

$$\mu^*(V) \leq \mu^*(A) + \varepsilon. \tag{A.13}$$

Since $V \cap U$ is open, there exists $f \in C(X)$ with $f \prec V \cap U$ and

$$\beta(f) \geq \mu^*(V \cap U) - \varepsilon.$$

Similarly, there exists $g \in C(X)$ with $g \prec V - \operatorname{supp}(f)$ and

$$\beta(g) \geq \mu^*(V - \operatorname{supp} f) - \varepsilon.$$

Then $f + g \prec V$, and so

$$\begin{aligned}\mu^*(V) &\geq \beta(f) + \beta(g)\\ &\geq \mu^*(V \cap U) + \mu^*(V - \operatorname{supp} f) - 2\varepsilon\\ &\geq \mu^*(A \cap U) + \mu^*(A - U) - 2\varepsilon.\end{aligned}$$

Using (A.13) and then taking $\varepsilon \to 0$ give

$$\mu^*(A) \geq \mu^*(A \cap U) + \mu^*(A - U).$$

Since the opposite inequality is automatic by subadditivity, this proves that U is measurable. Therefore, μ is a Borel measure on X.

The final step is to prove the integral formula (A.7), which will also establish the uniqueness of μ. Consider $f \in C(X)$ with $f \geq 0$. We first claim that

$$\mu\{f \geq 1\} \leq \beta(f) \leq \mu(\operatorname{supp} f). \tag{A.14}$$

The upper bound follows from the definition (A.8), because $f \prec U$ for any $U \supset \operatorname{supp}(f)$. For the lower bound, let $U_\varepsilon := \{f > 1 - \varepsilon\}$ for $\varepsilon > 0$. For $g \prec U_\varepsilon$ we have $f \geq (1 - \varepsilon)g$, which by the positivity of β implies that

$$\beta(g) \leq (1 - \varepsilon)^{-1}\beta(f).$$

Taking the supremum over all $g \prec U_\varepsilon$ gives

$$\mu(U_\varepsilon) \leq (1 - \varepsilon)^{-1}\beta(f).$$

Therefore, by (A.8),

$$\mu\{f \geq 1\} \leq (1 - \varepsilon)^{-1}\beta(f).$$

Taking $\varepsilon \to 0$ yields the lower bound in (A.14).

To refine the estimate (A.14), we fix some $n \in \mathbb{N}$ and decompose f into layers of height $1/n$ by setting

$$f_j(x) := \begin{cases} 0, & f(x) < j/n, \\ f(x) - j/n, & j/n \le f(x) \le (j+1)/n, \\ 1/n, & f(x) > (j+1)/n. \end{cases}$$

Note that $0 \le f_j \le 1/n$, and

$$f = \sum_{j=1}^{m} f_j \tag{A.15}$$

for m large enough that $\sup f \le m/n$.

If we set $K_j := \operatorname{supp}(f_j)$, then the fact that $\chi_{K_{j+1}} \le n f_j \le \chi_{K_j}$ implies

$$\mu(K_{j+1}) \le n \int_X f_j \, d\mu \le \mu(K_j),$$

by the monotonicity of the integral. By (A.15), we can sum over j to obtain

$$\frac{1}{n} \sum_{j=1}^{m} \mu(K_j) \le \int_X f \, d\mu \le \frac{1}{n} \sum_{j=0}^{m} \mu(K_j). \tag{A.16}$$

(Note that $K_m = \emptyset$.)

On the other hand, we can apply (A.14) to the function $n f_j$ to conclude that

$$\mu(K_{j+1}) \le n\beta(f_j) \le \mu(K_j).$$

Summing over j gives

$$\frac{1}{n} \sum_{j=1}^{m} \mu(K_j) \le \beta(f) \le \frac{1}{n} \sum_{j=0}^{m} \mu(K_j).$$

In conjunction with (A.16), this shows that

$$\left| \int_X f \, d\mu - \beta(f) \right| \le \frac{\mu(K_0)}{n}.$$

Taking $n \to \infty$ then completes the proof of (A.7) for $f \ge 0$. The general case follows by linearity. □

For the application of the Riesz representation theorem to spectral measures in Section 5.1.2, we need to know that $C(X)$ is dense in $L^2(X, d\mu)$. This holds for general Borel measures under certain regularity conditions. For measures obtained via the Riesz theorem, we can give a simple direct proof.

Lemma A.11. *Let X be a compact metric space and μ a Borel measure constructed from a positive functional on $C(X)$ as in Theorem A.10. Then $C(X)$ is dense in $L^p(X, d\mu)$ for $1 \leq p < \infty$.*

Proof By the definition of the integral on a measure space, simple functions are dense in L^p. It therefore suffices to show that the characteristic function χ_E for a Borel subset $E \subset X$ can be approximated in the L^p sense by continuous functions. Given $\varepsilon > 0$, the definition of μ^* gives an open set $U \supset E$ such that $\mu(U - E) < \varepsilon$. Then, by the definition of μ in terms of the functional β, there exists $f \in C(X)$ such that $f \prec U$ and

$$\mu(U) - \int_X f \, d\mu < \varepsilon.$$

We can then estimate

$$\begin{aligned} \|\chi_E - f\|_p &\leq \|\chi_E - \chi_U\|_p + \|\chi_U - f\|_p \\ &\leq 2\varepsilon^{\frac{1}{p}}. \end{aligned}$$

□

The Riesz representation theorem can be extended to more general topological spaces. The limiting factor is essentially the existence of the partition of unity. For a locally compact Hausdorff space, Urysohn's lemma implies the existence of locally finite partitions of unity, and the Riesz theorem can be extended to this case by a very similar argument. For further details, see, e.g., Rudin [78, Thm. 6.19] or Folland [31, §7.1].

A.2 L^p Spaces

Let $(X, \mathcal{M}, \mu)$ be a measure space. For $p \in [1, \infty)$, the L^p norm of a measurable function $f : X \to \mathbb{C}$ is defined by

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}. \tag{A.17}$$

For $p = \infty$, the integral is replaced by the essential supremum,

$$\|f\|_\infty := \inf\{m \in \mathbb{R} : \ |f| \leq m \text{ a.e.}\}.$$

The L^p spaces are defined as

$$L^p(X, d\mu) := \{f \text{ measurable } X \to \mathbb{C} : \ \|f\|_p < \infty\}, \tag{A.18}$$

subject to the standard equivalence of functions that agree almost everywhere with respect to μ.

The function $\|\cdot\|_p$ is homogeneous because of the power $1/p$ included in (A.17), and positive definiteness is a consequence of the equivalence relation imposed on $L^p(X, d\mu)$. The L^p version of the triangle inequality is known as the Minkowski inequality. It is obvious for $p = 1$ or ∞, and follows from Cauchy–Schwarz for $p = 2$. Its proof in the general case relies on the following:

Lemma A.12 (Hölder Inequality). *Let f, g be measurable functions on X. For $p, q \geq 1$ with $1/p + 1/q = 1$,*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q. \tag{A.19}$$

Proof Assume that p, q are as in the statement. For $x > 0$, calculus shows that the function

$$h(x) := x - \frac{x^p}{p}$$

is maximized when $x = 1$. Therefore,

$$x - \frac{x^p}{p} \leq 1 - \frac{1}{p}$$

for all $x > 0$. Setting $x = ab^{-q/p}$ for $a, b > 0$ gives, after some simplification, the inequality

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}, \tag{A.20}$$

which clearly extends to the case where a or $b = 0$,

Now suppose f, g are measurable functions on X. The inequality (A.19) is trivial if either $f \equiv 0$ or $g \equiv 0$, so we can assume that these functions have nonzero norms. Setting $a = |f(x)|/\|f\|_p$ and $b = |g(x)|/\|g\|_q$ in (A.20) gives

$$\frac{|f(x)g(x)|}{\|f\|_p \|g\|_q} \leq \frac{1}{p} \frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q} \frac{|g(x)|^q}{\|g\|_q^q}.$$

Integration over x gives

$$\frac{\|fg\|_1}{\|f\|_p \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1,$$

yielding (A.19). □

Corollary A.13 (Minkowski Inequality). *Let f, g be measurable functions on X. For $p \geq 1$,*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof As in the proof of Hölder, we can assume f and g are not $\equiv 0$, since this case is trivial. From (A.19), we have

$$\|f\|_p \geq \|fh\|_1 \tag{A.21}$$

under the assumption that $\|h\|_q = 1$ where $q = p/(p-1)$. Since equality holds in (A.21) for $h = |f|^{p-1}/\|f^{p-1}\|_q$, we can conclude that

$$\|f\|_p = \sup_{\|h\|_q=1} \|fh\|_1,$$

The triangle inequality for L^p now follows,

$$\begin{aligned} \|f+g\|_p &= \sup_{\|h\|_q=1} \|(f+g)h\|_1 \\ &\leq \sup_{\|h\|_q=1} \|fh\|_1 + \sup_{\|h\|_q=1} \|gh\|_1 \\ &= \|f\|_p + \|g\|_p. \end{aligned}$$

□

On $\mathbb{R}^n$, a *step function* is defined as a linear combination of characteristic functions of rectangles. From the construction of Lebesgue measure described in Appendix A.1.1, we can deduce that the step functions are dense in $L^p(\mathbb{R}^n)$ for $p \in [1,\infty)$. (See, e.g., Royden [76, §6.4].) By smoothing the edges of the characteristic functions, we can thus conclude that $C_0^\infty(\mathbb{R}^n)$ is also dense as a subset of $L^p(\mathbb{R}^n)$.

A.2.1 Completeness

This section is devoted to the proof of the completeness of L^p spaces (Theorem 2.5). The result is a straightforward consequence of the convergence theorems from the Lebesgue integration theory.

Theorem A.14 (Riesz–Fischer). *For $p \in [1,\infty]$, $L^p(X,d\mu)$ is complete as a metric space.*

Proof By Theorem 2.4, it suffices to show that an absolutely convergent series is convergent. Consider first the case $p = \infty$. For $\{u_k\}_{k=1}^\infty \subset L^\infty(X,d\mu)$, set $m_k := \|u_k\|_\infty$. We assume that the series is absolutely convergent, which means that

$$\sum m_k < \infty. \tag{A.22}$$

Since each $|u_k|$ is bounded by m_k almost everywhere, we can define an exceptional set

$$E := \bigcup_{k=1}^{\infty} \{x : |u_k(x)| > m_k\},$$

of measure zero, such that $|u_k| \leq m_k$ for all k on $X \backslash E$. For $x \notin E$, the series $\sum u_k(x)$ converges absolutely by (A.22). We can thus define a function $f \in L^\infty(X, d\mu)$ by

$$f(x) = \sum_{k=1}^{\infty} u_k(x), \quad \text{for } x \notin E,$$

with the values on E being irrelevant because the set has measure zero. The function f lies in $L^\infty(X, d\mu)$, with

$$\|f\|_\infty \leq \sum_{k=1}^{\infty} m_k.$$

It follows from (A.22) that $\sum u_k$ converges to f in the L^∞ sense, because

$$\left\| f - \sum_{k=1}^{n} u_k \right\|_\infty \leq \sum_{k=n+1}^{\infty} m_k.$$

Now let $p \in [1, \infty)$ and assume that $\sum u_k$ is an absolutely convergent series in $L^p(X, d\mu)$. Define

$$g(x) := \sum_{k=1}^{\infty} |u_k(x)|.$$

The triangle inequality implies that

$$\left\| \sum_{k=1}^{N} |u_k| \right\|_p \leq \sum_{k=1}^{N} \|u_k\|_p. \tag{A.23}$$

By the monotone convergence theorem, the left-hand side of (A.23) converges to $\|g\|_p$ as $N \to \infty$, implying that

$$\|g\|_p \leq \sum_{k=1}^{\infty} \|u_k\|_p.$$

Since $\sum u_k$ converges absolutely, this shows $g \in L^p(X, d\mu)$.

In particular, g is finite a.e., so the series $\sum u_k(x)$ converges absolutely for a.e. $x \in X$. Hence the series $\sum u_k$ converges pointwise a.e. to some function f. Moreover, $|f| \leq g$, so $f \in L^p(X, d\mu)$ also. Since

$$\left| \sum_{k=0}^{m} u_k - f \right|^p \leq (2g)^p,$$

and g^p is integrable, the dominated convergence theorem implies that

$$\lim_{m\to\infty} \int_X \left| \sum_{k=0}^{m} u_k - f \right|^p d\mu = 0.$$

Hence the series $\sum u_k$ converges to f in L^p. □

A.2.2 Convolution

The convolution of two measurable functions on $\mathbb{R}^n$ is defined by

$$f * g(x) := \int_{\mathbb{R}^n} f(x-y)g(y)\, d^n y, \tag{A.24}$$

assuming the integral is well defined. For $f, g \in L^1(\mathbb{R}^n)$, we can deduce from Fubini's theorem that the integral (A.24) exists for almost every x, defining a function $f * g \in L^1(\mathbb{R}^n)$ which satisfies

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1.$$

This basic result can be extended to combinations of L^p spaces, as follows.

Theorem A.15 (Young's Convolution Inequality). *Suppose $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$. Then, if r satisfies*

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}, \tag{A.25}$$

then $f * g \in L^r(\mathbb{R}^n)$ and

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

Proof It suffices to consider the case $f, g \in C_0^\infty(\mathbb{R}^n)$. Applying Hölder's inequality twice yields a triple product version,

$$\|fgh\|_1 \leq \|f\|_s \|g\|_t \|h\|_r, \tag{A.26}$$

where

$$\frac{1}{s} + \frac{1}{t} + \frac{1}{r} = 1.$$

To prove Young's inequality, we first divide up the convolution integrand into three terms,

$$\begin{aligned} |f * g(x)| &\le \int_{\mathbb{R}^n} |f(x-y)g(y)|\, d^n y \\ &\le \int_{\mathbb{R}^n} |f(x-y)|^{1-\frac{p}{r}} \cdot |g(y)|^{1-\frac{q}{r}} \cdot \left| f(x-y)^p g(y)^q \right|^{\frac{1}{r}} d^n y. \end{aligned}$$

Applying the Hölder inequality (A.26) to this expression gives

$$|f * g(x)| \le \left\| f^{1-\frac{p}{r}} \right\|_s \left\| g^{1-\frac{q}{r}} \right\|_t \left\| f(x-\cdot)^{\frac{p}{r}} g^{\frac{q}{r}} \right\|_r. \tag{A.27}$$

Assuming that (A.25) holds, we can choose s and t so that

$$\frac{1}{s} = \frac{1}{p} - \frac{1}{r}, \qquad \frac{1}{t} = \frac{1}{p} - \frac{1}{r}.$$

With these choices, (A.27) gives

$$|f * g(x)|^r \le \|f\|_p^{1-\frac{p}{r}} \|g\|_q^{1-\frac{q}{r}} \int_{\mathbb{R}^n} |f(x-y)|^p\, |g(y)|^q\, d^n y. \tag{A.28}$$

For the remaining integral over x, note that

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x-y)|^p\, |g(y)|^q\, d^n y\, d^n x \le \|f\|_p^p\, \|g\|_q^q\,,$$

by Fubini's theorem. Integrating (A.28) over x thus yields

$$\begin{aligned} \|f * g\|_r^r &\le \|f\|_p^{r-p} \|g\|_q^{r-q} \|f\|_p^p \|g\|_q^q \\ &= \|f\|_p^r \|g\|_q^r. \end{aligned}$$

□

A.3 Fourier Transform

In this section we review some standard background material on the Fourier transform, which is used extensively in the text.

The *Fourier transform* of a function in $f \in L^1(\mathbb{R}^n)$ is defined by

$$\hat{f}(\xi) := (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\, d^n x. \tag{A.29}$$

As a map the transform is denoted by

$$\mathcal{F} : f \mapsto \hat{f}.$$

The properties of the Fourier transform of an integrable function are characterized by the following:

Lemma A.16 (Riemann–Lebesgue). *For $f \in L^1(\mathbb{R}^n)$, the Fourier transform $\hat{f}$ is continuous and bounded, with*

$$\lim_{|\xi|\to\infty} \hat{f}(\xi) = 0.$$

Proof From (A.29) we see immediately that

$$|\hat{f}(\xi)| \le (2\pi)^{-\frac{n}{2}} \|f\|_1 \tag{A.30}$$

for all $\xi \in \mathbb{R}^n$. This also shows that $\mathcal{F}$ maps L^1 convergent sequences to uniformly convergence sequences.

For $f \in L^1(\mathbb{R}^n)$, let $\{\psi_k\} \subset C_0^\infty(\mathbb{R}^n)$ be an approximating sequence such that $\psi_k \to f$ in L^1. Using integration by parts, we can check that $\hat{\psi}_k$ is smooth and approaches zero as $|\xi| \to \infty$. Since $\hat{\psi}_k \to \hat{f}$ uniformly on $\mathbb{R}^n$, it follows that $\hat{f}$ is continuous and decays to zero at infinity. □

The primary goal of this section is to explain how the definition of $\mathcal{F}$ is extended to L^2 functions, for which the integral (A.29) may not exist.

Theorem A.17 (Plancherel). *The Fourier transform defined by* (A.29) *extends to a unitary map $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$.*

To prove Plancherel's theorem, we first study the restriction of $\mathcal{F}$ to the space of *Schwartz functions*,

$$\mathcal{S}(\mathbb{R}^n) := \left\{ f \in C^\infty(\mathbb{R}^n);\ \left\| x^\alpha D^\beta f \right\|_\infty < \infty \text{ for all } \alpha, \beta \in (\mathbb{N}_0)^n \right\}. \tag{A.31}$$

For a Schwartz function $\psi \in \mathcal{S}(\mathbb{R}^n)$, integration by parts implies that

$$\mathcal{F}[D_x^\alpha \psi](\xi) = (i\xi)^\alpha \hat{\psi}(\xi), \tag{A.32}$$

and

$$\mathcal{F}[x^\alpha \psi](\xi) = (iD_\xi)^\alpha \hat{\psi}(\xi). \tag{A.33}$$

This means that, under the Fourier transform, smoothness translates to rapid decay and vice versa.

Lemma A.18. *The Fourier transform $\mathcal{F}$ maps $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$.*

Proof Suppose that $f \in \mathcal{S}$. In order to show that $\hat{f}$ is Schwartz, we need to produce a bound on the function $\xi^\beta D^\alpha \hat{f}$ for each α, β. By (A.32) and (A.33),

$$\xi^\beta D_\xi^\alpha \hat{f}(\xi) = i^{|\alpha|+|\beta|} \int_{\mathbb{R}^n} e^{-i\xi\cdot x} x^\alpha D_x^\beta f(x)\, d^n x. \tag{A.34}$$

Because $(1+|x|^2)^{-N}$ is integrable for N sufficiently large, we can estimate (A.34) by

$$\left|\xi^\beta D_\xi^\alpha \hat{f}(\xi)\right| \le C_N \sup \left|(1+|x|^2)^N x^\alpha D_x^\beta f\right|.$$

The right side is finite for all N by the definition (A.31). □

The final ingredient for the proof of Theorem A.17 is the formula for the Fourier transform of a Gaussian function. Let

$$g(x) := e^{-a|x|^2}, \tag{A.35}$$

for $a > 0$. By completing the square, and then using contour integration to make a complex change of variables, we can compute

$$\hat{g}(\xi) = (2a)^{-\frac{n}{2}} e^{-|\xi|^2/4a}. \tag{A.36}$$

Theorem A.19. *The Fourier transform on $\mathcal{S}(\mathbb{R}^n)$ defined by (A.29) has an inverse $\mathcal{F}^{-1}$ given by*

$$f(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{i\xi\cdot x} \hat{f}(\xi)\, d^n\xi. \tag{A.37}$$

Proof For $f, g \in \mathcal{S}(\mathbb{R}^n)$, consider the integral

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x) e^{-ix\cdot y} g(y)\, d^n x\, d^n y. \tag{A.38}$$

The integrals over x and y can be taken in either order, by Fubini, yielding the identity

$$\int_{\mathbb{R}^n} f\hat{g}\, d^n x = \int_{\mathbb{R}^n} \hat{f} g\, d^n y. \tag{A.39}$$

Now let g be the Gaussian function (A.35). By (A.36),

$$(2a)^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(x) e^{-|x|^2/4a} \, d^n x = \int_{\mathbb{R}^n} \hat{f}(y) e^{-a|y|^2} \, d^n y. \tag{A.40}$$

Rescaling the variable on the left-hand side of (A.40) yields

$$2^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(\sqrt{a} x) e^{-|x|^2/4} \, d^n x = \int_{\mathbb{R}^n} \hat{f}(y) e^{-a|y|^2} \, d^n y.$$

By the dominated convergence theorem, taking $a \to 0$ then gives

$$(2\pi)^{\frac{n}{2}} f(0) = \int_{\mathbb{R}^n} \hat{f}(y) \, d^n y. \tag{A.41}$$

This verifies (A.37) for $x = 0$.

The general inverse formula can be deduced from (A.41) by a simple translation argument. For $w \in \mathbb{R}^n$, define the translation operator T_w on $\mathcal{S}(\mathbb{R}^n)$ by

$$T_w f(x) := f(x + w).$$

A change of variables shows that

$$\begin{aligned} \widehat{T_w f}(y) &= \int_{\mathbb{R}^n} e^{-ix \cdot y} f(x + w) \, d^n x \\ &= \int_{\mathbb{R}^n} e^{-i(x-w) \cdot y} f(x) \, d^n x \\ &= e^{iw \cdot y} \hat{f}(y). \end{aligned}$$

Plugging $T_w f$ into (A.41) yields

$$(2\pi)^{n/2} f(w) = \int_{\mathbb{R}^n} e^{iw \cdot y} \hat{f}(y) \, d^n y.$$

□

From the pairing formula (A.39) and the invertibility of $\mathcal{F}$, we can immediately deduce that

$$\langle \hat{f}, \hat{g} \rangle = \langle f, g \rangle \tag{A.42}$$

for $f, g \in \mathcal{S}(\mathbb{R}^n)$. It is straightforward to extend a unitary map from a dense subspace to the full Hilbert space, so Theorem A.17 follows from (A.42), Theorem A.19, and the fact that $\mathcal{S}(\mathbb{R}^n)$ is dense as a subspace of $L^2(\mathbb{R}^n)$.

A.4 Elliptic Regularity

In Section 6.3 we gave a simple argument for the interior regularity of eigenfunctions. This approach, based on the Fourier characterization of Sobolev spaces, is easily generalized to elliptic PDE with constant coefficients. To handle elliptic operators with variable coefficients, or to include regularity up to the boundary, a different strategy is required.

In this section we will prove the more general elliptic regularity result stated in the text as Theorem 9.27. For more general versions of this result, see, e.g., Evans [29, §6.3.2], Gilbarg and Trudinger [36, §8.4], or Taylor [89, §5.1]

Theorem A.20 (Elliptic Regularity). *Let $\overline{\Omega}$ be a compact Riemannian manifold with boundary, with $-\Delta$ be the Dirichlet Laplacian defined in Section 9.3. If $u \in \mathcal{D}(\Delta)$ and $\Delta u \in H^m(\Omega)$ for some $m \in \mathbb{N}_0$, then $u \in H^{m+2}(\Omega)$, with*

$$\|u\|_{H^{m+2}} \le C\big(\|\Delta u\|_{H^m} + \|u\|\big),$$

where C depends only on Ω and m.

The strategy for the proof of Theorem A.20 is to use difference quotients to estimate weak derivatives. For $h \in \mathbb{R}$ and f a function on $\mathbb{R}^n$, define

$$\partial_j^h f(x) := \frac{f(x+he_j) - f(x)}{h}$$

for $j = 1, \ldots n$, where $\{e_j\}$ denotes the standard basis for $\mathbb{R}^n$. We first need to prove a basic estimate for difference quotients in terms of derivatives.

Lemma A.21. *Suppose $u \in H_0^1(U)$ where $U \subset \mathbb{R}^n$ is a bounded open set. For $\varepsilon > 0$, let $U_\varepsilon := \{x \in U : d(x, \partial U) > \varepsilon\}$. If $\operatorname{supp}(u) \subset U_\varepsilon$ for $\varepsilon > 0$, then for $|h| < \varepsilon$,*

$$\|\partial_j^h u\|_{L^2} \le \|\partial_k u\|_{L^2}.$$

Proof For $\psi \in C_0^\infty(\Omega)$, the difference quotient can be expressed as an integral,

$$\partial_j^h \psi(x) = \frac{1}{h}\int_0^h \partial_j \psi(x + te_j)\, dt,$$

for h sufficiently small. Applying Cauchy–Schwarz then gives the estimate,

$$|\partial_j^h \psi(x)|^2 \le \frac{1}{h}\int_0^h \big|\partial_j \psi(x+te_j)\big|^2\, dt.$$

Integrating over x then gives

$$
\begin{aligned}
\|\partial_j^h \psi\|^2 &\le \frac{1}{h}\int_U \int_0^h \left|\partial_j \psi(x+te_j)\right|^2 dt\,dx \\
&= \frac{1}{h}\int_0^h \int_U \left|\partial_j \psi(x+te_j)\right|^2 dx\,dt \\
&= \|\partial_j \psi\|^2.
\end{aligned}
$$

The estimate can now be extended to a function $u \in H_0^1(U)$ with support in U_ε, by choosing an approximating sequence $\{\phi_k\} \subset C_0^\infty(\Omega)$ such that $\phi_k \to u$ in H^1. □

Proof of Theorem A.20 The first observation is that the proof can be localized to coordinate neighborhoods. Let $\{U_j\}_{j=1}^q$ and $\{\chi_j\}$ be the coordinate atlas and corresponding partition of unity used to define the Sobolev norms as in (9.39). A simple concavity argument shows that

$$
\|u\|_{H^m} \asymp \sum_{j=1}^q \|\chi_j u\|_{H^m},
$$

for each m, with constants that depend only on q. Furthermore,

$$
\begin{aligned}
\|\Delta(\chi_j u)\|_{H^m} &\le \|\chi_j \Delta u\|_{H^m} + \|[\Delta, \chi_j]u\|_{H^m} \\
&\le \|\chi_j \Delta u\|_{H^m} + C\|u\|_{H^{m+1}}.
\end{aligned}
$$

We will first show that

$$
\|\chi_j u\|_{H^2} \le C\big(\|\Delta(\chi_j u)\|_{L^2} + \|\chi_j u\|_{L^2}\big), \tag{A.43}
$$

for each j. We only need to consider the case of a boundary neighborhood U_j, as the interior estimate can be considered as a special case where the cutoff vanishes near the boundary.

To prove (A.43), we can specialize to the case of a bounded domain $U := \{|x| < R, x^n > 0\}$ in $\mathbb{R}^n$, with the metric g represented in coordinates as matrix g_{ij}. On $H_0^1(U)$ we define the sesquilinear form

$$
\begin{aligned}
Q[u, v] &= \int_U g(\nabla \bar{u}, \nabla v)\, dV \\
&:= \int_U g^{ij}(\partial_i \bar{u})(\partial_j v)\sqrt{g}\, d^n x.
\end{aligned}
$$

For simplicity, we will write this as

$$Q[u, v] = \int_U a^{ij} (\partial_i \bar{u})(\partial_j v)\, d^n x,$$

where $a^{ij} := g^{ij}\sqrt{g}$.

Recall the local formula (9.33) for the Laplacian,

$$\Delta = \frac{1}{\sqrt{g}} \partial_i \big(\sqrt{g} g^{ij} \partial_j\big). \tag{A.44}$$

By the definition of $\mathcal{D}(-\Delta)$ from Section 9.4.1, we have

$$Q[u, v] = \langle -\Delta u, v\rangle, \tag{A.45}$$

for $u \in \mathcal{D}(-\Delta)$ and $v \in H_0^1(U)$.

For the application to (A.43), we may assume that u is supported away from the boundary of U, except possibly at $x^n = 0$. Assuming the $v \in H_0^1(U)$ shares this property, we can apply ∂_k^h to v for $k = 1, \ldots, n-1$ and h sufficiently small. We then deduce from (A.45) that

$$Q[u, \partial_k^h v] = \langle -\Delta u, \partial_k^h v\rangle. \tag{A.46}$$

Writing out the left-hand side gives

$$Q[u, \partial_k^h v] = \int_U a^{ij} (\partial_i \bar{u})(\partial_j \partial_k^h v)\, d^n x.$$

By a linear change of variables, the difference quotient can be transferred from one term to the other (as if integrating by parts) to obtain

$$Q[u, \partial_k^h v] = -\int_U \partial_k^{-h} (a^{ij} \partial_i \bar{u})\, \partial_j v\, dx.$$

A simple product-rule computation shows that

$$\partial_k^{-h} (a^{ij} \partial_i \bar{u}) = a^{ij}(x - he_k) \partial_i (\partial_k^{-h} \bar{u}) + (\partial_k^{-h} a^{ij}) \partial_i \bar{u}.$$

(As above, $\{e_k\}$ denotes the standard basis for $\mathbb{R}^n$.) Applying these computations to (A.46) gives

$$\int_U a^{ij}(x - he_k) \partial_i (\partial_k^{-h} \bar{u})\, \partial_j v\, dx = \int_U \Big[(L\bar{u}) \partial_k^h v - (\partial_k^{-h} a^{ij}) \partial_i \bar{u}\, \partial_j v\Big]\, dx.$$

By Lemma A.21, the L^2 norm of $\partial_k^h v$ can be estimated by that of $\partial_k v$, for h sufficiently small. Thus Cauchy–Schwarz gives the estimate

$$\begin{aligned}\int_U a^{ij}(x - he_k)\partial_i(\partial_k^{-h}\bar{u})\,\partial_j v\,dx &\le C\big(\|\Delta u\| + \|\nabla u\|\big)\|\nabla v\| \\ &\le C\big(\|\Delta u\| + \|u\|_{H^1}\big)\|\nabla w\|.\end{aligned} \tag{A.47}$$

Here C is independent of h, because the coefficients $(\partial_k^{-h}a^{ij})$ are bounded uniformly for small h, by the mean value theorem.

Now let us set $v = \partial_k^{-h}u$. The fact that g^{ij} is smooth and positive definite on $\overline{U}$ implies

$$\int_U a^{ij}(x - he_k)\partial_i(\partial_k^{-h}\bar{u})\,\partial_j(\partial_k^{-h}u)\,dx \ge c\left\|\partial_k^{-h}\nabla u\right\|^2.$$

Applying this in (A.47) gives

$$\left\|\partial_k^{-h}\nabla u\right\|^2 \le C\big(\|\Delta u\| + \|u\|_{H^1}\big)\left\|\partial_k^{-h}\nabla u\right\|,$$

which yields

$$\left\|\partial_k^{-h}\nabla u\right\| \le C\big(\|\Delta u\| + \|u\|_{H^1}\big). \tag{A.48}$$

By Cauchy–Schwarz, we can estimate

$$\begin{aligned}\|u\|_{H^1}^2 &= \langle -\Delta u, u\rangle \\ &\le \|\Delta u\|\|u\| \\ &\le \frac{1}{2}(\|\Delta u\| + \|u\|)^2.\end{aligned}$$

Therefore, (A.48) can be reduced to

$$\left\|\partial_k^{-h}\nabla u\right\| \le C\big(\|\Delta u\| + \|u\|\big), \tag{A.49}$$

with C independent of h.

By Alaoglu's theorem (Theorem 2.37), the uniform estimate (A.49) implies that there exists functions $f_j \in L^2(U)$ satisfying

$$\|f_j\| \le C\big(\|\Delta u\| + \|u\|\big),$$

and a sequence $h_l \to 0$ such that $\partial_k^{-h_l}\partial_j u \to f_j$ weakly as $l \to 0$. For $\psi \in C_0^\infty(U)$ we have

$$\begin{aligned}\int_U f_j \psi \, d^n x &= \lim_{l\to 0} \int_U (\partial_k^{-h_l} \partial_j u) \psi \, d^n x \\ &= -\lim_{l\to 0} \int_U \partial_j u \, \partial_k^{h_l} \psi \, d^n x \\ &= -\int_U \partial_j u \, \partial_k \psi \, d^n x.\end{aligned}$$

This shows that $f_j = \partial_k \partial_j u$ as a weak derivative, and (A.49) gives the estimate

$$\|\partial_k \nabla u\| \leq C\big(\|\Delta u\| + \|u\|\big). \tag{A.50}$$

This argument required $k \leq n-1$, so the estimate (A.50) covers all of the second partial derivatives of u except for $\partial_n^2 u$. To handle this case, note that (A.44) implies that

$$\Delta u = \frac{1}{\sqrt{g}}\Bigg[\sum_{(i,j)\neq(n,n)} \partial_i [a^{ij} \partial_j u] + a^{nn} \partial_n^2 u + (\partial_n a^{nn}) \partial_n u\Bigg].$$

All of the terms here except $a^{nn} \partial_n^2 u$ are in $L^2(U)$, either by assumption or by (A.50). Since a^{nn} is bounded below by a positive constant, by the positive definiteness of the metric on $\overline{U}$, it follows that $\partial_n^2 u \in L^2(U)$, with

$$\left\|\partial_n^2 u\right\| \leq C\big(\|\Delta u\| + \|u\|\big).$$

Together with (A.50), this shows that $u \in H^2(U)$, with

$$\|u\|_{H^2} \leq C\big(\|\Delta u\| + \|u\|\big). \tag{A.51}$$

This completes the proof of (A.43), which settles the case $m = 0$.

To prove the estimate for higher m we proceed by induction. Assume that $u \in H^{m+1}(U)$ (with support as described above) and satisfies

$$\|u\|_{H^{m+1}} \leq C\big(\|\Delta u\|_{H^{m-1}} + \|u\|_{H^m}\big).$$

For the inductive step we assume that $\Delta u \in H^m(U)$, and need to show that this implies $u \in H^{m+2}(U)$, with the corresponding estimate.

Since $u \in H^{m+1}$, the weak derivative $D^\alpha u$ exists for $|\alpha| = m$ and lies in $H_0^1(\Omega)$. It satisfies the equation

$$\Delta(D^\alpha u) = D^\alpha(\Delta u) - [\Delta, D^\alpha] u. \tag{A.52}$$

By the assumptions on u and Δu, and the fact that $[\Delta, D^\alpha]$ is a differential operator of order $m + 1$, the right-hand side of (A.52) lies in $L^2(\Omega)$. We can therefore apply the bound (A.51) to $D^\alpha u$ to conclude that $D^\alpha u \in H^2(\Omega)$ with

$$\|D^\alpha u\|_{H^2} \leq C\big(\|\Delta(D^\alpha u)\| + \|D^\alpha u\|_{H^1}\big).$$

Furthermore, by (A.52) we can estimate

$$\|\Delta(D^\alpha f)\| \leq \|\Delta u\|_{H^m} + C\|u\|_{H^{m+1}}.$$

It thus follows that

$$\|u\|_{H^{m+2}} \leq C\big(\|\Delta u\|_{H^m} + \|u\|_{H^{m+1}}\big).$$

□

References

1. Arendt, W., Nittka, R., Peter, W., Steiner, F.: Weyl's law: spectral properties of the Laplacian in mathematics and physics. In: Mathematical Analysis of Evolution, Information, and Complexity, pp. 1–71. Wiley-VCH, Weinheim (2009)
2. Aronszajn, N.: A unique continuation theorem for solutions of elliptic partial differential equations or inequalities of second order. J. Math. Pures Appl. (9) **36**, 235–249 (1957)
3. Arveson, W.: A Short Course on Spectral Theory. Graduate Texts in Mathematics, vol. 209. Springer, Berlin (2002)
4. Ashbaugh, M.S.: Isoperimetric and universal inequalities for eigenvalues. In: Spectral Theory and Geometry (Edinburgh, 1998). London Mathematical Society Lecture Notes Series, vol. 273, pp. 95–139. Cambridge University Press, Cambridge (1999)
5. Ashbaugh, M.S., Benguria, R.D.: Proof of the Payne-Pólya-Weinberger conjecture. Bull. Am. Math. Soc. **25**, 19–29 (1991)
6. Avila, A., Jitomirskaya, S.: The Ten Martini Problem. Ann. Math. (2) **170**, 303–342 (2009)
7. Baker, M., Rumely, R.: Harmonic analysis on metrized graphs. Can. J. Math. **59**, 225–275 (2007)
8. Berger, M.: A Panoramic View of Riemannian Geometry. Springer, Berlin (2003)
9. Berger, M., Gauduchon, P., Mazet, E.: Le spectre d'une variété Riemannienne. Lecture Notes in Mathematics, vol. 194. Springer, Berlin (1971)
10. Berkolaiko, G.: An elementary introduction to quantum graphs. In: Geometric and Computational Spectral Theory. Contemporary Mathematics, vol. 700, pp. 41–72. American Mathematical Society, Providence (2017)
11. Berkolaiko, G., Kuchment, P.: Introduction to Quantum Graphs. Mathematical Surveys and Monographs, vol. 186. American Mathematical Society, Providence (2013)
12. Birkhoff, G., Kreyszig, E.: The establishment of functional analysis. Hist. Math. **11**, 258–321 (1984)
13. Borthwick, D.: Introduction to Partial Differential Equations. Universitext. Springer, Cham (2016)
14. Borthwick, D.: Spectral Theory of Infinite-Area Hyperbolic Surfaces. Progress in Mathematics, vol. 318, 2nd edn. Birkhäuser/Springer, Cham (2016)
15. Brooks, R.: A relation between growth and the spectrum of the Laplacian. Math. Z. **178**, 501–508 (1981)
16. Brooks, R.: Inverse spectral geometry. In: Progress in Inverse Spectral Geometry. Trends in Mathematics, pp. 115–132. Birkhäuser, Basel (1997)
17. Buser, P.: Geometry and Spectra of Compact Riemann Surfaces. Birkhäuser, Boston (1992)

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1

18. Carleman, T.: Propriétés asymptotiques des fonctions fondamentales des membranes vibrantes. In: Åttonde Skandinaviska Matematikerkongressen (Stockholm, 1934), pp. 34–44. Håkan Ohlssons Boktryckeri, Lund (1935)
19. Chavel, I.: Eigenvalues in Riemannian Geometry. Academic, London (1984). Including a chapter by Randol, B, With an appendix by Dodziuk, J.
20. Chavel, I.: Isoperimetric Inequalities. Cambridge Tracts in Mathematics, vol. 145. Cambridge University Press, Cambridge (2001)
21. Chernoff, P.R.: Essential self-adjointness of powers of generators of hyperbolic equations. J. Funct. Anal. **12**, 401–414 (1973)
22. Chung, F.R.K.: Spectral Graph Theory. CBMS Regional Conference Series in Mathematics, vol. 92. American Mathematical Society, Providence (1997)
23. Davies, E.B.: Heat Kernels and Spectral Theory. Cambridge Tracts in Mathematics, vol. 92. Cambridge University Press, Cambridge (1989)
24. Davies, E.B.: Spectral Theory and Differential Operators. Cambridge Studies in Advanced Mathematics, vol. 42. Cambridge University Press, Cambridge (1995)
25. Dieudonné, J.: History of Functional Analysis. North-Holland Mathematics Studies, vol. 49. North-Holland, Amsterdam (1981). Notas de Matemática [Mathematical Notes], p. 77
26. do Carmo, M.P.A.: Riemannian Geometry. Mathematics: Theory and Applications. Birkhäuser, Boston (1992). Translated from the second Portuguese edition by Francis Flaherty
27. Donnelly, H., Li, P.: Pure point spectrum and negative curvature for noncompact manifolds. Duke Math. J. **46**, 497–503 (1979)
28. Edmunds, D.E., Evans, W.D.: Spectral Theory and Differential Operators. Oxford Mathematical Monographs, Oxford University Press, Oxford (2018)
29. Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, 2nd edn., vol. 19. American Mathematical Society, Providence (2010)
30. Federer, H.: Geometric Measure Theory. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer, New York (1969)
31. Folland, G.B.: Real Analysis: Modern Techniques and Their Applications. Wiley, London (1984)
32. Friedlander, L.: Extremal properties of eigenvalues for a metric graph. Ann. Inst. Fourier (Grenoble) **55**, 199–211 (2005)
33. Gaffney, M.P.: The harmonic operator for exterior differential forms. Proc. Nat. Acad. Sci. U.S.A. **37**, 48–50 (1951)
34. Gårding, L.: On the asymptotic distribution of the eigenvalues and eigenfunctions of elliptic differential operators. Math. Scand. **1**, 237–255 (1953)
35. Gelfand, I.M.: Normierte Ringe. Rec. Math. [Mat. Sbornik] N. S. **9**(51), 3–24 (1941)
36. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Grundlehren der Mathematischen Wissenschaften, vol. 224, 2nd edn. Springer, Berlin (1983)
37. Gohberg, I.C., Krein, M.: Introduction to the Theory of Linear Nonselfadjoint Operators. Translations of Mathematical Monographs, vol. 18. American Mathematical Society, Providence (1969)
38. Gordon, C., Webb, D., Wolpert, S.: Isospectral plane domains and surfaces via Riemannian orbifolds. Invent. Math. **110**, 1–22 (1992)
39. Guillemin, V., Pollack, A.:Differential Topology. Prentice-Hall, Englewood Cliffs (1974)
40. Gustafson, S.J., Sigal, I.M.: Mathematical Concepts of Quantum Mechanics. Universitext, 2nd edn. Springer, Heidelberg (2011)
41. Hall, B.C.: Quantum Theory for Mathematicians. Graduate Texts in Mathematics, vol. 267. Springer, New York (2013)
42. Harper, P.G.: Single band motion of conduction electrons in a uniform magnetic field. Proc. Phys. Soc. Lond. A **68**, 874–878 (1955)
43. Heinonen, J.: Lectures on Lipschitz analysis. Report. University of Jyväskylä Department of Mathematics and Statistics, vol. 100. University of Jyväskylä, Jyväskylä (2005)
44. Hislop, P.D., Sigal, I.M.: Introduction to Spectral Theory. Applied Mathematical Sciences, vol. 113. Springer, Berlin (1996). With applications to Schrödinger operators

45. Hofstadter, D.R.: Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields. Phys. Rev. B **14**, 2239–2249 (1976)
46. Ivrii, V.: 100 years of Weyl's law. Bull. Math. Sci. **6**, 379–452 (2016)
47. Iwaniec, H.: Spectral Methods of Automorphic Forms. Graduate Studies in Mathematics, vol. 53, 2nd edn. American Mathematical Society, Providence (2002)
48. Kac, M.: Can one hear the shape of a drum? Am. Math. Mon. **73**, 1–23 (1966)
49. Kato, T.: Perturbation Theory for Linear Operators. Springer, Berlin (1995). Reprint of the 1980 edition
50. Kato, T.: Fundamental properties of Hamiltonian operators of Schrödinger type. Trans. Am. Math. Soc. **70**, 195–211 (1951)
51. Kato, T.: Perturbation of continuous spectra by trace class operators. Proc. Jpn Acad. **33**, 260–264 (1957)
52. Klingenberg, W.P.A.: Riemannian Geometry. De Gruyter Studies in Mathematics, vol. 1, 2nd edn. Walter de Gruyter, Berlin (1995)
53. Kuchment, P.: Quantum graphs: an introduction and a brief survey. In: Proceedings of Symposia in Pure Mathematics, Analysis on Graphs and Its Applications, vol. 77, pp. 291–312. American Mathematical Society, Providence (2008)
54. Kurasov, P., Naboko, S.: Rayleigh estimates for differential operators on graphs. J. Spectr. Theory **4**, 211–219 (2014)
55. Lablée, O.: Spectral Theory in Riemannian Geometry. EMS Textbooks in Mathematics. European Mathematical Society (EMS), Zürich (2015)
56. Lamé, G.: Mémoire sur la propagation de la chaleur dans les polyèdres. J. Ecol. Polytech. **22**, 194–251 (1833)
57. Lee, J.M.: Riemannian Manifolds. An Introduction to Curvature. Graduate Texts in Mathematics, vol. 176. Springer, New York (1997)
58. Lee, J.M.: Introduction to Smooth Manifolds. Graduate Texts in Mathematics, vol. 218. Springer, New York (2003)
59. Lieb, E.H., Loss, M.: Analysis. Graduate Studies in Mathematics, vol. 14, 2nd edn. American Mathematical Society, Providence (2001)
60. MacCluer, B.D.: Elementary Functional Analysis. Graduate Texts in Mathematics, vol. 253. Springer, New York (2009)
61. Magnus, W., Winkler, S.: Hill's Equation. Dover, New York (1979). Corrected reprint of the 1966 edition
62. Minakshisundaram, S., Pleijel, A.: Some properties of the eigenfunctions of the Laplace-operator on Riemannian manifolds. Can. J. Math. **1**, 242–256 (1949)
63. Monna, A.F.: Functional Analysis in Historical Perspective. Wiley, New York (1973)
64. Olver, F.W.J., Olde Daalhuis, A.B., Lozier, D.W., Schneider, B.I., Boisvert, R.F., Clark, C.W., Miller, B.R., Saunders, B.V.: NIST Digital Library of Mathematical Functions (2016). http://dlmf.nist.gov/. Release 1.0.19
65. Payne, L.E., Pólya, G., Weinberger, H.F.: On the ratio of consecutive eigenvalues. J. Math. Phys. **35**, 289–298 (1956)
66. Petersen, P.: Riemannian Geometry. Graduate Texts in Mathematics, vol. 171, 3rd edn. Springer, Berlin (2016)
67. Pleijel, Å.: A study of certain Green's functions with applications in the theory of vibrating membranes. Ark. Mat. **2**, 553–569 (1954)
68. Pólya, G., Szegő, G.: Isoperimetric Inequalities in Mathematical Physics. Annals of Mathematics Studies, vol. 27. Princeton University Press, Princeton (1951)
69. Reed, M., Simon, B.: Methods of Modern Mathematical Physics. I. Functional Analysis. Academic, London (1972)
70. Reed, M., Simon, B.: Methods of Modern Mathematical Physics. II. Fourier Analysis, Self-adjointness. Academic, London (1975)
71. Reed, M., Simon, B.: Methods of Modern Mathematical Physics. IV. Analysis of Operators. Academic, London (1978)

72. Reed, M., Simon, B.: Methods of Modern Mathematical Physics. III. Scattering Theory. Academic, London (1979)
73. Rellich, F.: Störungstheorie der Spektralzerlegung. Math. Ann. **116**, 555–570 (1939)
74. Riesz, F., Sz.-Nagy, B.: Functional Analysis. Frederick Ungar Publishing, New York (1955). Translated by Leo F. Boron
75. Roelcke, W.: Über den Laplace-Operator auf Riemannschen Mannigfaltigkeiten mit diskontinuierlichen Gruppen. Math. Nachr. **21**, 131–149 (1960)
76. Royden, H.L.: Real Analysis, 3rd edn. Macmillan, New York (1988)
77. Rudin, W.: Principles of Mathematical Analysis. International Series in Pure and Applied Mathematics, 3rd edn. McGraw-Hill, New York (1976)
78. Rudin, W.: Real and Complex Analysis, 3rd edn. McGraw-Hill, New York (1987)
79. Rudin, W.: Functional Analysis. International Series in Pure and Applied Mathematics, 2nd edn. McGraw-Hill, New York (1991)
80. Schmüdgen, K.: Unbounded Self-adjoint Operators on Hilbert Space. Graduate Texts in Mathematics, vol. 265. Springer, Dordrecht (2012)
81. Schoen, R., Yau, S.-T.: Lectures on Differential Geometry. In: Proceedings of the Conference on Lecture Notes in Geometry and Topology, I. International Press, Cambridge (1994)
82. Simon, B.: Semiclassical analysis of low lying eigenvalues. I. Nondegenerate minima: asymptotic expansions. Ann. Inst. H. Poincaré Sect. A (N.S.) **38**, 295–308 (1983)
83. Simon, B.: Operator Theory. A Comprehensive Course in Analysis, Part 4. American Mathematical Society, Providence (2015)
84. Sobolev, S.L.: Applications of Functional Analysis in Mathematical Physics. Translated from the Russian by F. E. Browder. Translations of Mathematical Monographs, vol. 7. American Mathematical Society, Providence (1963)
85. Sokal, A.D.: A really simple elementary proof of the uniform boundedness theorem. Am. Math. Mon. **118**, 450–452 (2011)
86. Steen, L.A.: Highlights in the history of spectral theory. Am. Math. Mon. **80**, 359–381 (1973)
87. Stein, E.M., Shakarchi, R.: Real Analysis. Princeton Lectures in Analysis, III. Princeton University Press, Princeton (2005)
88. Szegő, G.: Inequalities for certain eigenvalues of a membrane of given area. J. Rational Mech. Anal. **3**, 343–356 (1954)
89. Taylor, M.E.: Partial Differential Equations. I. Basic Theory. Springer, New York (1996)
90. Taylor, M.E.: Partial Differential Equations. II. Qualitative Studies of Linear Equations. Springer, New York (1996)
91. Teschl, G.: Ordinary Differential Equations and Dynamical Systems. Graduate Studies in Mathematics, vol. 140. American Mathematical Society, Providence (2012)
92. Venkov, A.B.: Spectral Theory of Automorphic Functions and Its Applications. Kluwer Academic Publishers, Dordrecht (1990)
93. von Neumann, J.: Charakterisierung des Spektrums eines Integralesoperators. Actualitës Sci. Ind. **229**, 38–55 (1935)
94. Weidmann, J.: Linear Operators in Hilbert Spaces. Graduate Texts in Mathematics, vol. 68. Springer, New York (1980). Translated from the German by Joseph Szücs
95. Weinberger, H.F.: An isoperimetric inequality for the N-dimensional free membrane problem. J. Rational Mech. Anal. **5**, 633–636 (1956)
96. Yafaev, D.R.: Mathematical Scattering Theory. Analytic Theory. Mathematical Surveys and Monographs, vol. 158. American Mathematical Society, Providence (2010)
97. Zworski, M.: Semiclassical Analysis. Graduate Studies in Mathematics, vol. 138. American Mathematical Society, Providence (2012)

Index

D. Borthwick, *Spectral Theory*, Graduate Texts in Mathematics 284,
https://doi.org/10.1007/978-3-030-38002-1